VOLUME 19, 1957/58, No. 1 pp. 1-40 Published 27th July 1957

Philips Technical Review

DEALING WITH TECHNICAL PROBLEMS
RELATING TO THE PRODUCTS, PROCESSES AND INVESTIGATIONS OF
THE PHILIPS INDUSTRIES

ELECTROLUMINESCENCE AND IMAGE INTENSIFICATION

by G. DIEMER, H. A. KLASENS and P. ZALM.

535.376:621.383.2:
621.3.032.36

The phenomenon of electroluminescence, which was discovered in 1936, has so far not fulfilled the expectations it aroused as regards its usefulness as a light source: although it can now be exploited in "luminous panels" producing light with reasonable efficiency (up to about 10 lumens/watt), it seems that such panels, supplied from the mains, would have to have very large surface areas to give the luminous flux normally required for lighting a room. On the other hand, the fact that local differences of brightness can be produced in a luminous panel in a simple fashion has opened up unexpected perspectives. On this fact are based the solid-state image intensifiers that are arousing increasing interest.

Introduction

An ideal light source would be one which converted all the energy supplied to it (e.g. heat or electrical energy) into light of the desired spectral composition (e.g. that of daylight) in the simplest possible way. Up to now the ideal is most nearly approached by the fluorescent lamp, a combination of the gas discharge and the phenomenon of fluorescence. There nevertheless remains a demand for a technically simpler process of energy conversion, one for example which would do away with the need for a vacuum-tight envelope (as is required by a gas discharge) and which would not involve corrosion or any other kind of ageing of the light source. The discovery of the phenomenon of electroluminescence [1]) by the Frenchman Destriau in the 'thirties does seem to have been a step forward on the road to direct conversion of electrical energy into light without ageing of the source, although we must admit that, despite recent advances in this field, light sources based on this phenomenon still do not seem to be of practical use for general lighting.

A light source based on electroluminescence is simple enough: it merely consists of say, a layer of suitably prepared substance (usually zinc sulphide), coated on a transparent base, and placed in an electric field. At voltages considerably lower than the breakdown voltage of the layer, the passage of current through the zinc sulphide results in the emission of light, no physical or chemical changes arising in the substance, and the heat developed is barely perceptible. An electroluminescent layer of this kind is conveniently named a *luminous panel.*

The photometric quantity used for measuring the light output of a source such as this, which radiates from an extended area, is the luminous emittance, this being the total number of lumens radiated from one side of the plane per square metre of its area. Luminous panels made so far have only had a moderate emittance (1 m^2 of an electroluminescent panel may at most emit as many lumens as a 200 W incandescent lamp) and not particularly high efficiency; for the time being, therefore, the employment of the luminous panel as a light source is likely to be confined to such special cases as the lighting of instrument dials, radio receiver scales, clock faces, etc. However, the fact that local differences of emittance can be produced in a simple way is of great interest, and attempts are being made to exploit this in solid-stage image intensifiers [2]) to which we shall devote a word at the end of this article.

[1]) G. Destriau, J. Chim. phys. **33**, 587-625, 1936; Phil. Mag. (7) **38**, 700-739, 774-793, 880-888, 1947. For a survey of the subject of electroluminescence, especially that of zinc sulphide powders, see P. Zalm, Philips Res. Rep. **11**, 353-399 and 417-451, 1956 (Nos. 5 and 6); G. Destriau and H. F. Ivey, Proc. Inst. Rad. Engrs., Solid-state materials issue, December 1955.

[2]) See for example G. Diemer, H. A. Klasens and J. G. van Santen, Philips Res. Rep. **10**, 401-424, 1955.

The article will be mainly concerned with a description of the complicated mechanism of light production in a luminous panel. We shall first discuss one or two important properties of the device.

Construction and general behaviour of a luminous panel

A diagram showing the construction of a luminous panel appears in *fig. 1*. The panel consists of a glass base to which a transparent electrode is applied,

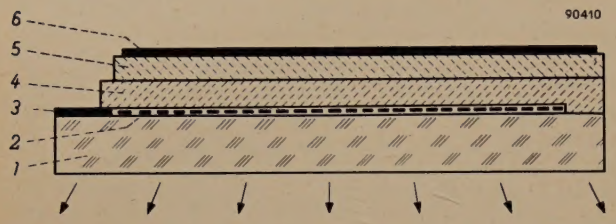


Fig. 1. Diagram showing construction of an electroluminescent cell or "luminous panel". *1* glass plate (base). *2* very thin, transparent, conducting coating of tin oxide. *3* metal contact strip. *4* layer of zinc sulphide suspended in a transparent lacquer (approx. 25 microns thick). *5* reflector layer of white lacquer (approx. 20 microns thick). *6* metallic layer, constituting the second electrode. Between this last and the tin oxide coating is applied an alternating voltage of several hundred volts at a frequency in the audio range (50-10 000 c/s). A second glass plate is usually glued on to electrode *6* to protect the panel from damage due to ingress of moisture.

e.g. a very thin layer of tin oxide sprayed on at high temperature. On top of the transparent electrode is applied a layer of electroluminescent substance, consisting of a powder (zinc sulphide) prepared in a special way and suspended in a transparent insulating lacquer. This second layer is about 25 microns thick. In order that the light produced in it should as far as possible be reflected

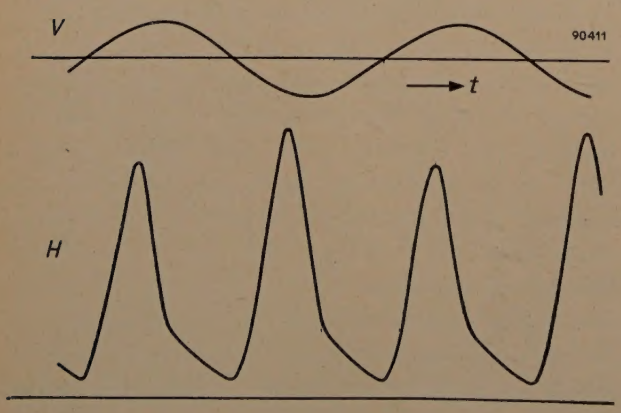


Fig. 2. Sinusoidal alternating voltage V and the electroluminescent effect (emittance H) it produces, as functions of time. A flash is produced every half-period. The frequency of the voltage depicted here was 65 c/s.

outwards via the transparent electrode, it is generally given a coating of white lacquer before the panel is backed with a second (metallic) electrode. Thus the panel as a whole represents a parallel plate condenser.

If we apply a sufficiently high alternating voltage to the panel, a flash of light is seen to be emitted during each half-period of oscillation (*fig. 2*). Sometimes, depending on the nature of the electroluminescent substance and the frequency of the voltage applied, a second and less intense flash is observed. Provided the frequency is neither too

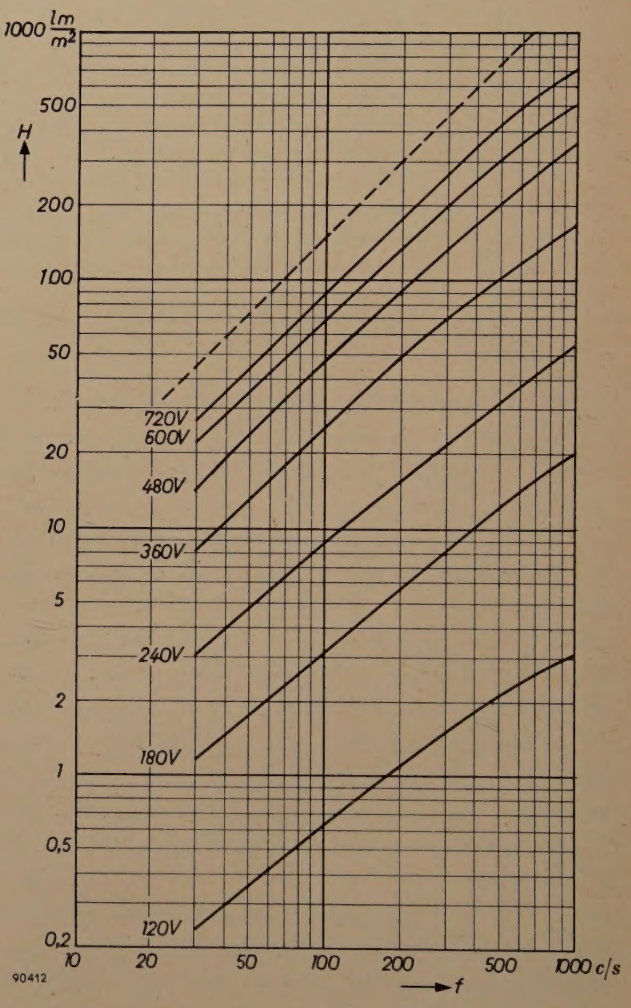


Fig. 3. If the r.m.s. value of voltage V applied to the cell be kept constant, emittance H increases near enough directly proportional to frequency f. (The curves are roughly parallel to the broken line at 45°.)

low (not below 10 c/s) nor too high (not above 1000 c/s), the amount of light emitted in each flash is found to be more or less independent of it; hence H, the emittance of the panel averaged out over a period of time, becomes greater for the same voltage at a higher frequency (being roughly in direct proportion to it). This can be seen from *fig. 3*.

With frequency kept constant, changes in the voltage V are found to produce very large changes in emittance. At low values of voltage the emittance increases sharply with increasing voltage, less sharply at higher values. A simple formula relating the emittance H and the voltage V is found by plotting the logarithm of the emittance (which is quite a good measure for our subjective impression of brightness) against the reciprocal of the root of the voltage (see *fig. 4*). This gives a straight line with a negative slope:

$$\ln H = \ln H_0 - \frac{\text{const.}}{\sqrt{V}} \quad . \quad . \quad . \quad . \quad (1)$$

In order therefore to obtain a high emittance the panel must be operated at a high frequency and at a voltage as high as its breakdown strength will permit.

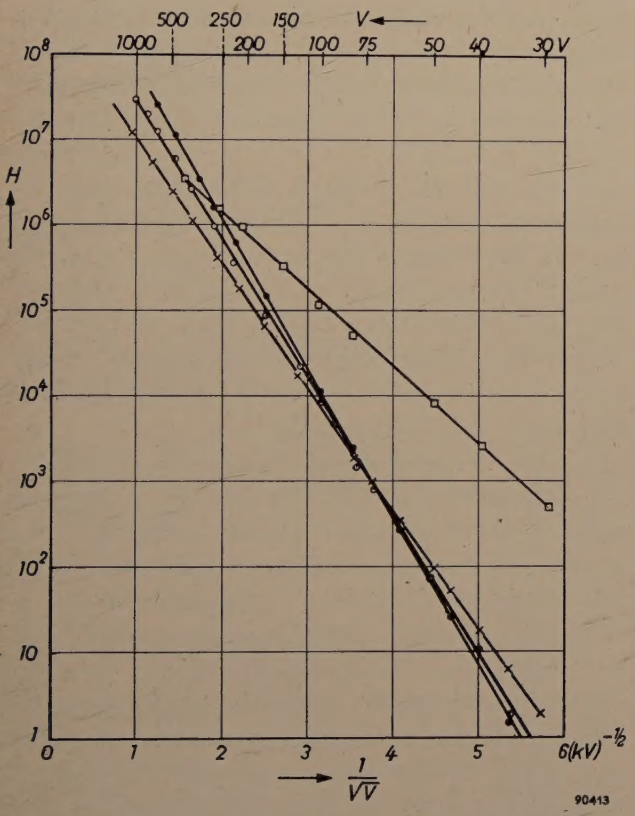


Fig. 4. The logarithm of the emittance H is a linear function of $1/\sqrt{V}$. The above curves were obtained experimentally, with various zinc sulphide powders, the H scale being in arbitrary units.

Mechanism of electroluminescence

The functioning of a luminous panel differs from other familiar kinds of luminescence, such as cathodoluminescence (as in television picture tubes) and photoluminescence (as in fluorescent lamps), only in the way in which the excitation of the luminescent substance takes place [3]). In cathodoluminescence, excitation is brought about by bombardment with a beam of fast electrons which, in penetrating the substance, raise certain bound electrons to a higher energy state; the bound electrons are thereby rendered capable of emitting light quanta $h\nu$ (h = Planck's constant = 6.62×10^{-34} joule sec; ν = frequency of the light = number of vibrations per second) on reverting to the ground state. In photoluminescence the same kind of excitation is caused to take place in the substance by its absorption of short-wave radiation, such as ultra-violet quanta, falling upon it. It seems that electroluminescence, as in a luminous panel, can best be understood by supposing that mobile electrons are present in the electroluminescent substance, that these electrons are able to acquire energy from the electric field and that, when they have sufficient energy, are able to bring bound electrons in the crystal into an excited state by colliding with them (impact excitation).

We shall now discuss the mechanism of excitation by impact in somewhat more detail.

Let us take ZnS as an example of an electroluminescent substance. A crystal of this substance, taken as a whole, is electrically neutral; electrical neutrality will also be found to obtain within each small element of its volume. Let us suppose the crystal to be coated at opposite ends with electrodes and that a strong electric field F is set up in it by a voltage on the electrodes. If now by some means a few electrons are introduced into the ZnS crystal, they will move freely in the direction of the field, as they would in a gas discharge, travelling an average distance of, say, l. In this way they are accelerated by the field up to an energy of eFl before giving up their acquired energy in collisions (e = charge of the electron = 1.6×10^{-19} coulomb). In an ideal crystal with its atoms or ions arranged with complete regularity, the mean free path l may amount to many interatomic distances, despite the density of the solid. (This means that the chance of collision is very much less than in a gas of the same density, where there is no regular arrangement.) In ZnS, for example, $l \approx 100$ Å, this being roughly equivalent to 40 interatomic distances. Collision being a matter of chance, there will be some electrons that travel a much longer distance than 100 Å before they collide. The theory of probability gives the following relation for n, the

[3]) For a general introductory discussion of fluorescent substances (phosphors) and the mechanism of fluorescence, see for example F. A. Kröger, Philips tech. Rev. **6**, 349-358, 1941.

number of electrons having a free path of x instead of the mean free path l:

$$\ln n = \ln n_0 - \frac{x}{l}, \quad \ldots\ldots \quad (2)$$

where n_0 is a constant proportional to the total number of electrons. Let the energy necessary for the excitation process described above be E_v. To acquire this amount of energy the electron must travel a distance x, without meeting an obstacle, such that

$$eFx = E_v,$$

i.e.

$$x = \frac{E_v}{eF}. \quad \ldots\ldots \quad (3)$$

Inserting this in (2), we obtain the following for the number of successful collisions:

$$\ln n_x = \ln n_0 - \frac{E_v}{eFl}. \quad \ldots\ldots \quad (4)$$

If we assume that there is a certain fixed chance of the bound electron giving rise to the emission of light on reverting to the ground state, then in the stationary state the rate of excitation must be proportional to the rate of emission. We can therefore compare the excitation formula, eq. (4), with eq. (1), the formula for the emittance found experimentally. On doing so, we see that both portray the same kind of dependence on voltage provided that the field strength F that accelerates the electrons is proportional to the *root* of the applied voltage. This seems strange at first sight because, after all, the field strength in a plate condenser (in which the field is homogeneous) is nothing other than the applied voltage divided by the distance between the electrodes, and is therefore proportional to the voltage itself, and not to its root.

We must seek a connection between this apparent inconsistency and another remarkable fact: perceptible emission of light begins at voltages far below the breakdown voltage of pure zinc sulphide (at voltages 10 to 100 times lower, in fact), whereas it is only when the breakdown voltage has been reached that the field strength — as calculated for a homogeneous field — has increased to the point where the high-energy collisions necessary for excitation become really numerous [4]). This paradox can be resolved by assuming that the field in the crystal is *not* homogeneous and that a considerable part of the voltage applied to it acts across a very thin layer, wherein it produces a local field strength much higher than the average. In this very thin layer the electrons would have a much greater chance of collecting sufficient energy for excitation in the time elapsing before their next collision.

A "barrier" of this kind could indeed arise, on the cathode side of the crystal, say, in the following way. We are assuming that in the process of excitation electrons are liberated in the crystal, being freed by collision, so that they are able to move in the direction of the anode. If an electron is removed from an electrically neutral volume element, an equal positive charge will be left over. By adding certain foreign ions (in a concentration of 10^{-3} or 10^{-4}) having a lower valency than the zinc ions belonging to the lattice — copper ions, for example — the positive charge thus formed will remain bound in the vicinity of a nearby copper ion. At the same time, the excitation energy is stored in these bound positive charges: emission of light takes place when, on reversal of the field, the displaced electron recombines with the positive charge. If at one place in the crystal a number of positive charges be created (and the chance of this happening is greatest near the cathode, as we shall see later), then the field strength ceases to be constant throughout the crystal; the region of the positive space charge has a much greater density of lines of force than the average, because these lines of force all run from the positive space charge to the negative charge induced on the electrode nearby (see *fig.* 5). In this high-field region a high rate of excitation ensues in accordance with eq. (4), further increasing the space charge there and causing the field to contract ever more closely into a thin layer.

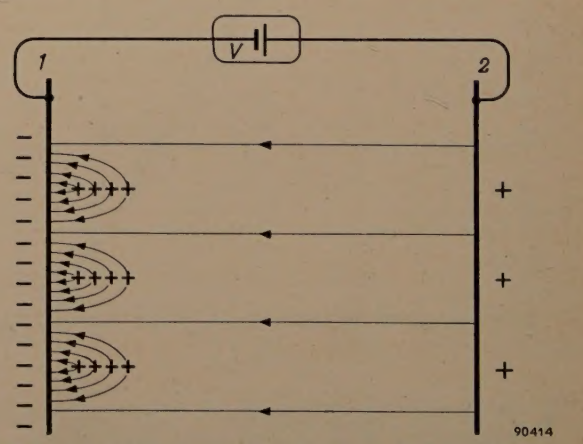


Fig. 5. In a zinc sulphide crystal provided with electrodes *1* and *2* connected to a voltage source V, a layer of positive space charge of certain thickness is set up on the cathode side of the crystal. The field strength in this layer, as indicated by the lines of force in the diagram, is much greater than the average field strength in the crystal.

[4]) This follows from eq. (4) if E_v be given the value 2.5 eV, i.e. the energy corresponding to a visible light quantum.

This explains the remarkable fact that emission of light occurs even at low voltages; it also allows us to explain why the field strength F, which according to eq. (4) is responsible for excitation, is proportional not to V, the applied voltage, but to its root. We assume that the accumulation of positive charges near immobile copper ions continues to the point where almost the whole of the voltage is acting across the space-charge layer and that, in conditions of maximum excitation, all the copper ions possess extra positive charges. If a higher voltage is applied to the crystal, there will be an increase in the total space charge Q under conditions of maximum excitation; but since all the copper ions already have excess charges, the increase in Q must be accompanied by a thickening of the space-charge layer, the depth of which in fact is proportional to Q. If we regard the space-charge layer as a charged capactitor with a plate separation of d (d being the thickness of the layer adjusted by a factor), we find that its capacitance C_b, which is proportional to $1/d$, decreases with increasing voltage. From the formula giving the charge on a capacitor:

$$Q = C_b \times V,$$

it follows that

$$V \propto d^2. \quad \ldots\ldots\ldots \quad (5)$$

However, F, the field strength in the barrier, is given by V/d, hence

$$F \propto d. \quad \ldots\ldots\ldots \quad (6)$$

Comparison of eqs. (5) and (6) shows that the field strength producing excitation is proportional to the root of the applied voltage, this bringing eq. (4) into line with eq. (1), the emittance law based on experimental observations.

Grains of zinc-sulphide powder suspended in a lacquer make no electrical contact with the electrodes and with one another. They therefore have to be provided with a very thin conductive skin or, to be more exact, with little conductive patches (to prevent short-circuiting from one side of the grain to the other). These "local electrodes" serve to collect the induced negative charge which, together with the positive space charge in the crystal, creates the strong field that causes excitation (see fig. 5). Moreover, even in the absence of an external voltage, there is a discontinuity in potential near the surface of the grain ("contact potential") due to the mere contact of the ZnS crystal and its conductive skin. This makes it understandable that excitation should tend to start in the surface of the grain on its cathode side, and that the subsequent formation of a positive space charge should cause the process to be gradually intensified until the stationary state is reached.

The conductive patches on the ZnS grains may be formed by adding to the zinc sulphide more copper than is soluble in it at a certain temperature; the remainder then separates out as a conductive compound of copper (probably copper sulphide). The presence of conductive patches on the surface of its grains is one of the most important characteristics distinguishing electroluminescent zinc sulphide from the normal luminescent kind.

As already stated, when an alternating voltage is applied to the panel the energy of excitation is released as light each time the field is reversed, the electrons being driven back to the positive charges and reverting to their ground state of energy ("recombination"). If an alternating voltage is suddenly applied at full strength to a virgin crystal, no light is emitted in the first half-period of oscillation. During the subsequent periods the recombination flash gradually becomes brighter until the stationary value is attained. This indicates that it is not possible during the first half-period to build up the space-charge barrier to the depth it ultimately reaches. If after some time the alternating voltage is withdrawn, and re-applied after a pause, the barrier builds up much more quickly than it does in a virgin crystal, even if the waiting interval has lasted several seconds; furthermore, a flash is now emitted during the first half-period. This proves that the recombination process taking place as the displaced electrons are swept back is not a complete one: the greater part of the positive space charge is left over each time, and that part recombines only very slowly after the field has been removed.

That light originates only in very small regions of the crystal is shown by microscopic examination. For a more extensive description and explanation of these and other details of the very complicated phenomenon of electroluminescence (the above-mentioned second flash in each period, for example), the reader is referred to the literature cited in footnote [1]).

So far we have been concerned solely with A.C. luminescence. Zinc sulphide phosphors also exist that emit light when a direct voltage is applied to them (provided arrangements are made for direct current to pass through the zinc sulphide). The emission of light is made possible by an activator which, when excited by impact, does not lose its electron but keeps it bound in a higher energy state.

In this way the electron is able to revert to the ground state, occasioning the emission of light, without the field being reversed. Manganese is an activator of this kind; in addition to manganese, however, it is necessary to introduce activators such as Cu into the ZnS in order to ensure that a space-charge barrier shall be formed.

The question might be posed as to where the first electrons come from that enable the excitation process to make a start, thereby allowing the barrier to build up. The answer cannot be given with certainty, but it is probable that these electrons come from the "local electrode" patches on the zinc sulphide. It is also possible that bound electrons occur in irregularities in the outermost atomic layers of the crystal itself and that these electrons can be released by the field acting in that region.

To conclude this explanation of the mechanism of electroluminescence it may be mentioned that a further light phenomenon exists which is frequently referred to as electroluminescence, and in which electrical energy in a solid is directly converted into light without first being converted into heat, but which is based on an entirely different mechanism. This is the phenomenon in which light is produced by the recombination of holes and electrons at a *p-n* junction [5]) when a voltage is applied to it in the forward direction. This kind of electroluminescence is found, for example, in SiC [6]), Ge [7]) and CdTe [8]). In a sense, it can be regarded as a reversal of the familiar and important photo-electric effect exhibited by *p-n* junctions. We shall not go further into the matter here.

Fabrication of the luminous panel

The foregoing will probably have made it clear that in every case an electroluminescent substance must also have fluorescent properties, in other words that it must be a phosphor [3]). Amongst all the many known phosphors, ZnS is eminently suitable as the basic material for electroluminescent purposes and there are two reasons for this. Firstly ZnS can be activated in such a way that electrons have a reasonably high mobility within it and are therefore capable of being accelerated up to sufficiently high energies. Secondly, ZnS so activated allows the space-charge barriers described above to build up easily. A further important point is that, in zinc sulphide, the Cu centres to which the positive charge is bound are stable in strong electric fields.

The conditions necessary for the chemical preparation of electroluminescent ZnS have been investigated in detail by Zalm [1]), making use of the earlier work of Kröger *et al.* [9]) on ZnS phosphors activated with copper. The solubility of Cu in ZnS is to a great extent influenced by the "coactivator" (Cl, Al, etc.); it is of importance that the concentration of Cu (expressed in gram-atoms) should be somewhat greater than that of the coactivator, to ensure that there is an excess of copper that can precipitate on cooling. Further, the concentrations of both are made sufficiently high to guarantee effective activation. As in other types of luminescence of ZnS activated with Cu, it is found that the product can be made to luminesce with a blue or green light according to the concentrations of activator and coactivator used. By further adding manganese an orange-yellow light can be obtained. By replacing sulphur by selenium in increasing proportions (that is, by taking mixed crystals of ZnS-ZnSe as basis for the powder), the light emitted by green or blue electroluminescent ZnS powders can be shifted into a continuous range of longer wavelengths. The addition of organic fluorescent substances capable of converting visible light of short wavelength into orange or red light makes it possible to produce panels giving a light of practically any desired colour.

A suspension of the electroluminescent powder is made in a synthetic lacquer which forms the binder; then, by means of known techniques such as spraying or "silk-screen printing", a homogeneous coating of the suspension about 35 μ thick is applied to a glass base previously rendered conducting by means of a tin-oxide coating. The lacquer should preferably have the following properties: 1) a high dielectric constant ε, to ensure that as much as possible of the alternating voltage acts across the ZnS grains (which themselves have a high ε); 2) a high breakdown voltage, so that the panel can be made to give high emittance values; and 3) a low moisture content (and also as moisture-*proof* as possible). Electrolysis may take place in an electroluminescent layer containing moisture, and this will shorten the life of the panel; moreover, moisture causes high dielectric losses that are non-productive of light. Urea formaldehyde resin was found to give a reasonably good compromise to these somewhat conflicting requirements.

---

[5]) See for example F. H. Stieltjes and L. J. Tummers, Philips tech. Rev. **17**, 233-246, 1955/56.

[6]) O. W. Lossew, Phys. Z. **34**, 397-403, 1933; C.R. Acad. Sci. U.R.S.S. **39**, 363-369, 1940.
K. Lehovec, C. A. Accardo and E. Jamgochian, Phys. Rev. **89**, 20-25, 1953.

[7]) J. R. Haynes and H. B. Briggs, Phys. Rev. **86**, 647, 1952.

[8]) C. Z. van Doorn and D. de Nobel, Physica **22**, 338-342, 1956 (No. 4).

[9]) F. A. Kröger, J. E. Hellingman and N. W. Smit, Physica **15**, 990-1018, 1949. N. W. Smit, Physica **16**, 317-328, 1950. F. A. Kröger and J. Dikhoff, Physica **16**, 297-316, 1950.

Nevertheless, a suspension of ZnS in a binder of this kind generally proves to have a fairly low breakdown voltage, because the ZnS grains with their conductive patches may easily form a bridge from one electrode to the other. Now the reflector layer of white lacquer (TiO_2 suspended in the same binder), which is sandwiched between the ZnS layer and the metallic electrode at the back, has a double function: besides reflecting the light emitted in the ZnS (towards the transparent electrode), it effects a considerable rise in the overall breakdown voltage. However, the reflector layer is not made unduly thick (not more than 20 μ), because of losses and the voltage drop across it which adversely affect the efficiency and emittance of the panel. Owing to the high dielectric constant of TiO_2, this drop is only about 25% of the overall voltage, but the reflector layer roughly doubles the breakdown voltage of the panel.

These layers are applied and baked-on successively. The whole is then backed with a metallic electrode by spraying with a colloidal suspension of silver. The panel is then thoroughly dried and, without delay, provided with a further backing in the shape of a glass plate, which is glued on to it; this is to prevent moisture from the atmosphere penetrating the somewhat hygroscopic layer of lacquer. Traces of moisture seriously diminish the life of the panel. If the layer is properly dried and protected, the life of the panel can extend to several thousands of hours, its emittance normally suffering a loss of about 10% within a short time (50 hours) of its being taken into use.

Good protection against the penetration of moisture can also be obtained by suspending the ZnS grains in a glaze instead of a lacquer; the suspension can be applied to a metallic base. However, in the present article we shall confine ourselves to electroluminescent layers having an organic binder.

The method of fabrication described above allows panels to be made in any desired shape and size, and this is of importance where the panel is to be used for instrument scale illumination, for example. It is equally easy to shape the panel into letters, figures and the like.

Emittance and efficiency

The flux Φ in lumens given by a luminous panel, is equal to the product of the emittance H and the area A of the panel. Another point of interest is its efficiency η, i.e. the ratio of flux Φ to the power P consumed (i.e. the number of lumens per watt).

The power dissipated by a lossy capacitor — the luminous panel is electrically equivalent to such a capacitor — can most easily be derived from the power factor $\tan\delta$, which is the ratio of the resistive component of current through the capacitor to the purely capacitive (wattless) current. With a voltage of r.m.s. value V across the panel, the power dissipated is:

$$P = V^2 \times 2\pi f C \tan\delta, \quad \ldots \quad (7)$$

and its efficiency is:

$$\eta = \frac{\Phi}{P} = \frac{HA}{V^2 \times 2\pi f C \tan\delta}, \quad \ldots \quad (8)$$

where f is the frequency of the alternating voltage and C is the capacitance of the panel.

Fig. 6 shows curves of H, $\tan\delta$ and η as functions of voltage, as derived from measurements on a certain electroluminescent layer. Curves showing how the same quantities vary with frequency appear in *fig. 7*.

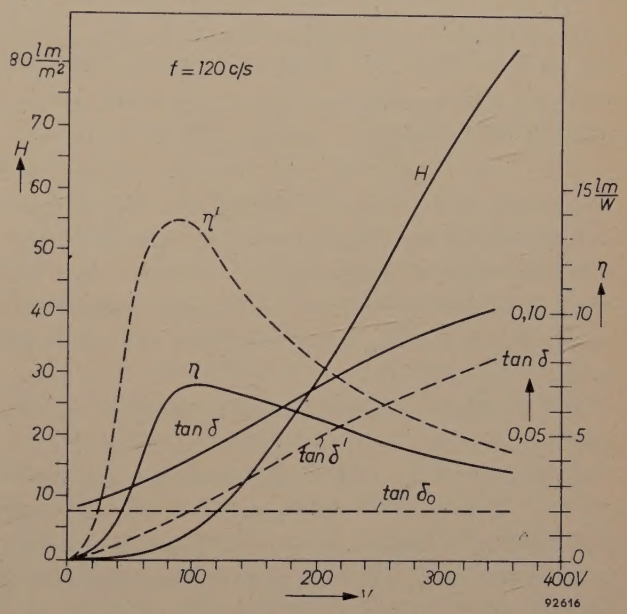


Fig. 6. Power factor $\tan\delta$, emittance H (in lumens/m²) and efficiency η (in lumens/watt) as functions of voltage V, the frequency being constant at $f = 120$ c/s, as obtained from measurements on a luminous panel giving green light. The dotted-line curves refer to the intrinsic efficiency η' (efficiency when the loss corresponding to $\tan\delta_0$ is neglected, see text).

Tan δ, the power factor, can be analyzed into two parts: a part $\tan\delta_0$, independent of voltage and proceeding from losses that are not directly connected with the mechanism of excitation (for the most part losses in the binder and in the series resistance of the tin oxide); and another part, $\tan\delta'$, which increases with voltage. This second part is connected with the power required for the transport of electrons through the ZnS crystals and for excitation. At higher voltages the resistive component of the

current corresponding to tan δ' increases more rapidly than the capacitive component (the rise of which with V is almost linear, since C is virtually independent of V); this indicates that the number of electrons injected increases with V at a more than linear rate. (Accordingly, we should, strictly speaking, introduce a correction into eq. (4) before comparing it with eq. (1); however, it can be shown that this correction does not essentially alter the voltage-dependence of the excitation as expressed by eq. (4) — see the article by Zalm cited in footnote [1]).)

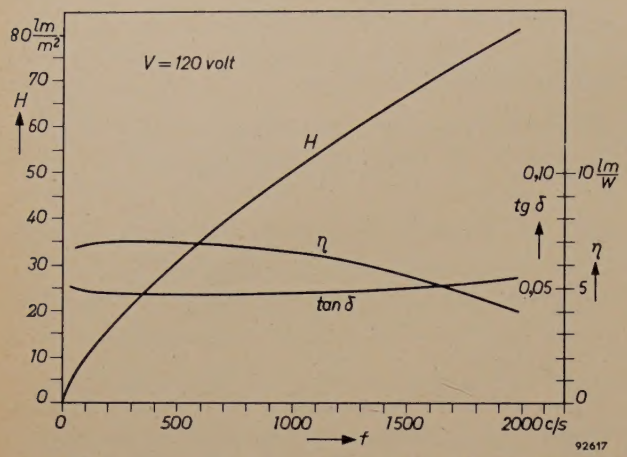


Fig. 7. The same quantities as in fig. 6 plotted as functions of f, the voltage being constant at $V = 120$ volts.

The maximum occurring in the curve of η as a function of V may be understood as follows. As a consequence of the way in which H depends on V (eq. 1), Φ increases rapidly with V at the start (at $V \approx 0$ it increases as an infinitely high power of V); as $V \to \infty$, however, the flux Φ approaches a limiting value, should electrical breakdown not already have taken place. Initially, the ratio $\Phi/P = \eta$ will likewise rapidly increase with V (as long as tan δ does not do so too), and it will attain a maximum at the point where numerator and denominator increase as the same power of V — that is, approximately at the point where Φ increases as the square of V. In reality, the maximum of η comes at a somewhat lower value of V, owing to the increase in tan δ with increasing V. For most electroluminescent layers, the quadratic increase of Φ with V occurs at a voltage between 100 and 200 volts.

As regards frequency dependence, it should be remarked that both C and tan δ are more or less independent of f. In electroluminescent layers with a large surface area, however, tan δ increases with frequency in consequence of the series resistance of the tin oxide layer. The measurements on which the present account is based were carried out on panels having an area of a few square centimetres only, in order to eliminate this complication. It will be clear why tan δ is independent of frequency if it be borne in mind that approximately equal charges are transported in each oscillatory period. Hence both the capacitive and resistive components of current increase more or less proportionally with frequency. Since Φ is approximately proportional to f, the efficiency η is more or less independent of f. The increase of Φ with f starts to become markedly less than linear at frequencies above 1000 c/s; the reason is that at these higher frequencies an ever smaller proportion of the electrons in the space-charge region has an opportunity to recombine with positive charges during the brief time that the field is of opposite sign. It will be seen that in oscillograms like that of fig. 2 the flashes begin to overlap considerably at these frequencies.

We may therefore conclude that maximum efficiency is obtained at moderate values of V, and hence for moderate values of emittance, and that at higher values of V higher values of H will be attained at the cost of lower efficiency. At higher frequencies H starts by increasing (linearly with f, roughly speaking) without η falling off. But in the kilocycle range η clearly begins to fall off; in panels with large surface areas the fall-off of η begins at still lower frequencies, because in such panels the series resistance of the tin oxide layer plays a bigger part.

Performance of existent electroluminescent layers

In recent years a considerable improvement in the emittance, efficiency and life of luminous panels has resulted from research [10]) into the optimum dimensioning of the panels as regards layer thickness, the filling fraction of ZnS in the binder, methods of applying the layers, and so on. In *Table I* some characteristic values are summarized for two panels having green emission (with and without an anti-breakdown TiO_2 reflector layer), and one panel having blue emission. The table shows the useful effect of the white TiO_2 layer in raising the breakdown voltage V_{cr}, the maximum efficiency η_m and the maximum emittance H_m attainable at low voltages. The blue panel has a much lower efficiency; this is due in part to the low sensitivity of the eye to blue light,

[10]) In the programme carried out in the Eindhoven laboratories, H. J. M. Joormann was responsible for the work on lacquers; measurements were carried out by T. J. Westerhof and K. W. C. Lugtenborg.

Table I. Data on ZnS luminous panels produced with urea-formaldehyde binder and having an area of about 7 cm². Each of the values given in the table is the average of those found for four panels made in the same way. In addition to the breakdown voltage V_{cr} and the relative dielectric constant ε, the table gives: the maximum efficiency (overall efficiency η_m, intrinsic efficiency η_m'), which is obtained at a voltage V_m; the emittance H_m obtained at the same voltage at a frequency of 50 c/s; emittance and efficiency at the maximum voltage permissible ($V = 0.6\ V_{cr}$) and a frequency of 50 c/s; and finally, emittance and efficiency at this maximum voltage and a frequency of 2000 c/s.

Composition of panel and thickness of layers	Colour of light emit.	V_{cr} (V)	ε	$V = V_m$ and $f = 50$ c/s			$V = 0.6\ V_{cr}$ and $f = 50$ c/s		$V = 0.6\ V_{cr}$ and $f = 2000$ c/s	
				η_m (lm/W)	η_m' (lm/W)	H_m (lm/m²)	H (lm/m²)	η (lm/W)	H (lm/m²)	η (lm/W)
45 μ ZnS	green	300	11	4.6	12	4	40	2	1000	1
29 μ ZnS + 23 μ TiO_2	green	550	12.8	10.2	17	4.6	70	4.5	1750	2
25 μ ZnS + 20 μ TiO_2	blue	400	9.8	0.8	—	0.3	3	—	90	—

in consequence of which fewer lumens are represented in a blue source than in a green one radiating the same amount of energy. Another reason for the low efficiency of the blue panel is that in the blue grains there is a higher power loss (higher tan δ). The η_m' column gives efficiency figures in the calculation of which only tan δ', the voltage-dependent part of the power factor, has been taken into account (for which, in other words, a fraction tan δ_0, caused by losses in the binder etc. having no connection with the mechanism of excitation, has been subtracted from the total power factor tan δ). It will be seen that even in the panel with the highest efficiency there is still a "useless" loss of 40%. If it were possible to eliminate this loss without reducing ε, the efficiency of the panel could be improved by a good 50%, thus becoming somewhat higher than that of an incandescent lamp. The efficiency of this hypothetical panel would approximate to the maximum efficiency theoretically obtainable from electroluminescence, as estimated by Zalm [1]).

The emittance of a panel connected to the 50 c/s mains, possibly via a transformer, can attain a value of 70 lm/m², i.e. about a tenth of the emittance of the screen of a television picture tube. Emittance can be greatly increased by connecting the panel to a generator of higher frequencies; supplied with A.C. at 2000 c/s, the best panel has an emittance of 1750 lm/m² and an efficiency roughly half of that which it has at 50 c/s (providing its surface area is not too large).

From these data it will be clear that for lighting a living room, for example, a panel area of the order of 1 square metre will certainly be necessary to obtain an adequate total luminous flux, and that the panel will have to be supplied at a frequency of some kilocycles, necessitating a special generator. The total efficiency of an installation like this, even for a panel emitting green light, will still be far below that of a normal incandescent lamp. This fact renders electroluminescence of little interest for normal lighting purposes. On the other hand, it offers an elegant solution to problems of scale illumination and the like: for this purpose there is no need for emittance and efficiency to be very high, the green colour is no objection, there is the advantage of easy shaping, and a further advantage deserving mention is the excellent uniformity of the brightness over the surface of a large panel.

However, the most interesting field of application of electroluminescent layers is undoubtedly their employment for intensifying images. To conclude this article, we shall give a short account of the "solid-state image intensifiers" based thereon. For the sake of brevity, and in order to distinguish these devices from image intensifiers of the vacuum-tube type [11]), we would propose for them the generic name of *amplificons* [12]).

Employment of electroluminescent layers for image intensification

There are two different methods of employing electroluminescent layers for this purpose. Both of them are, of course, based on a "control action" (as indeed are all forms of amplification); in other words, the energy flux is made to vary under the influence of a signal representing a very small, sometimes negligible energy. The first method can be referred to as that using external control, and the second as that using internal control.

---

[11]) M. C. Teves and T. Tol, Electronic intensification of fluoroscopic images, Philips tech. Rev. **14**, 33-43, 1952/53.

[12]) With regard to amplificons, see the article cited in footnote [2]) and also the December 1955 number of Proc. Inst. Rad. Engrs. (Solid-state materials issue).

Method using external control

Use is made of at least two layers, *I* and *II*, electrically connected in series (see *fig. 8*). Layer *II* is a normal electroluminescent layer having one electrode only, which is transparent. On top is applied a layer (*I*) of a photoconductive substance such as cadmium sulphide; *I* is also provided with one electrode which is transparent to the radiation (of intensity H_1) to be intensified. Voltages of V_1 and V_2 are set up across *I* and *II* respectively by applying a voltage $V = V_1 + V_2$ across the whole.

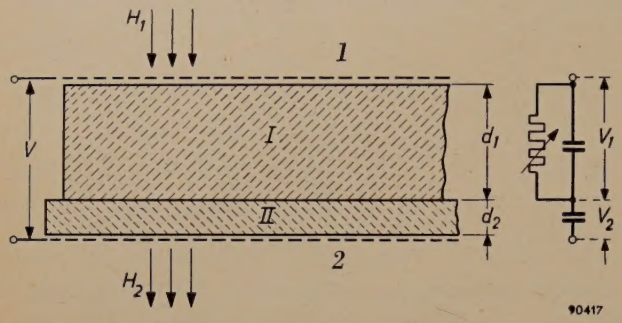


Fig. 8. Schematic diagram showing construction of an amplificon (solid-state image intensifier) using external control; the equivalent circuit is shown on the right. *1* and *2* transparent electrodes. *I* photoconductive layer of thickness d_1, across which acts a voltage V_1. *II* electroluminescent layer of thickness d_2, across which acts a voltage V_2. The two layers are electrically connected in series, and a voltage V ($= V_1 + V_2$) is applied across them. The radiation H_1 to be intensified falls on layer *I*; by virtue of electroluminescence in layer *II* an emittance H_2 is obtained.

In the absence of radiation a photoconductive substance possesses but few mobile electrons; when radiation containing quanta of sufficient energy falls upon it, however, some of its electrons are rendered mobile. Zinc sulphide too, to some extent, exhibits the phenomenon of photoconductivity (on this is based the functioning of the intensifier using internal control — see below), but cadmium sulphide does so to a far greater extent.

Now, the ratio between d_1 and d_2, the thicknesses of the two layers, can be made such that in the absence of incident radiation (that is, when $H_1 = 0$) almost the whole of the voltage applied acts across layer *I* (so that $V_2 \approx 0$ and $V_1 \approx V$). In these circumstances layer *II* emits little or no light (its emittance $H_2 \approx 0$). For this to be so, d_1 must be at least a few tenths of a millimetre thick for a d_2 of 50 microns.

If layer *I* is locally illuminated, the electrical resistance at the illuminated spots will diminish, and the greater the intensity H_1 of the incident radiation, the lower will be the resistance of the places affected. The result is that a smaller fraction of the applied voltage *V* acts across layer *I* at these places, V_2 becomes higher there and so does H_2. Thus layer *II* reproduces the local (and temporal) variations in H_1, and it does so in *intensified* measure: amplificons have already been made in which up to 50 quanta are emitted (H_2) for every quantum of incident radiation (H_1). The overall effect of the amplificon is described by the curve given by the relation between the local value of H_1 and the corresponding value of H_2. In *fig. 9* this curve has been constructed by combining the V_2-H_1 and H_2-V_2 characteristics, the latter being the "electroluminescence characteristic".

It will be observed that the relation between H_1 and H_2 is not linear, and this means that generally the gradation in the incident image is not faithfully reproduced. In principle, this makes it possible to heighten contrast, and the capacity to do so can be of importance in the intensification of X-ray images. However, by adopting special measures, such as the use of an alternating voltage containing two or more frequencies, the characteristic can be considerably straightened out.

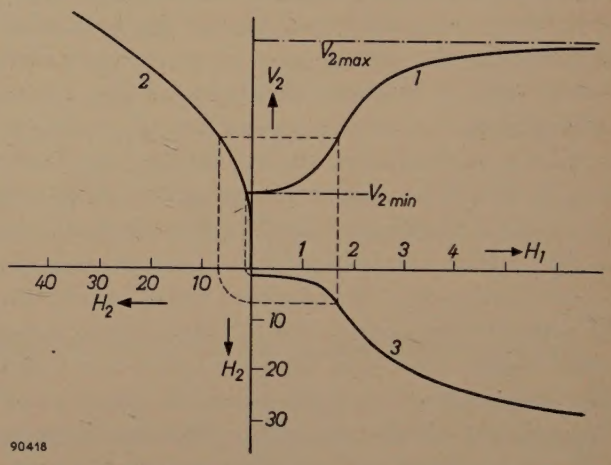


Fig. 9. Amplificon characteristic (*3*) constructed by combining the V_2-H_1 characteristic (*1*) and the H_2-V_2 or electroluminescence characteristic (*2*). $V_{2\,max}$ = maximum voltage acting across layer *II* when the radiation falling on layer *I* is very intense (when $H_1 \rightarrow \infty$). $V_{2\,min}$ = minimum voltage acting across layer *II* in the absence of incident radiation (when $H_1 = 0$).

Method using internal control

Destriau [13]) discovered that certain zinc sulphides activated with copper and manganese, which exhibit but little electroluminescence even at high voltages, might be caused by irradiation with ultra-violet or X-rays to emit light more strongly. In this case too, electroluminescence reproduces local and temporal variations in the

[13]) See p. 1911 of the article by Destriau and Ivey referred to in footnote [1]).

incident radiation, giving "intensification" of this radiation. The functioning of this kind of intensifier is to be understood as follows. The ultra-violet quanta absorbed by the intensifier render additional electrons mobile in the ZnS, thus furthering the mechanism described above by which space-charge layers are formed. In this way, the controlling of the process of electroluminescence takes place within the ZnS itself. In an image intensifier constructed on this principle by Cusano [14]), up to 10 visible quanta are emitted by the panel for every quantum of ultra-violet radiation it absorbs. At high levels of illumination the intensification effect falls off. The presence of manganese in the layer allows the intensifier to operate on D.C. as well as A.C.

For a more detailed discussion of amplificon design and performance the reader is referred to the literature as indicated in the footnotes. Although these types of image intensifier are still in their infancy and it is impossible to say with certainty how long it will take for them to evolve into a product suitable for manufacture, it is nevertheless very probable that they will find many applications in the future. It is conceivable that radar reception, fluoroscopy and perhaps television will be amongst them. For the last-mentioned, the inertia of the effect (namely, of the photoconduction process in the intensifier using external control), which still constitutes an obstacle at the present time, would have to be overcome.

[14]) D. A. Cusano, Phys. Rev. **98**, 546-547, 1955.

Summary. Layers of specially-prepared zinc sulphide powder, activated with copper, emit light when an electric alternating tension is applied to them. This effect, which is called electroluminescence and which was discovered by Destriau in 1936, can be applied to dial illumination in instruments, radio receivers and clocks and in general wherever uniformly luminous surfaces are required and where high efficiency and high luminous flux per unit area (emittance) are not essential. This method of light production does not promise much as far as general lighting purposes are concerned: at best (in layers emitting green light) its efficiency is just about equal to that of the incandescent lamp and, in order to raise emittance to a value high enough for the adequate lighting of a living room (for example) with a luminous surface of not more than, say, a square metre in area, the electrical supply has to be at a fairly high frequency, e.g. 2000 c/s, and the efficiency is then considerably reduced. In this article the more important details of the complicated phenomenon of electroluminescence, such as the way emittance and efficiency depend on voltage and frequency, are explained in terms of a mechanism whose validity is supported by a number of experimental facts. The essential feature of this mechanism is the formation, near the surface of each zinc-sulphide grain, of thin layers of strong positive space charge, where the field strength is much higher than the average; by virtue of the strong field, mobile electrons in the layers are capable of bringing about intense excitation by impact and, each time the field reverses, a certain proportion of the bound positive charges in the layers emit light on recombining with electrons. Finally, a brief account is given of the employment of the phenomenon in solid-state image intensifiers (amplificons), which may well find important applications in radar, X-ray fluoroscopy, etc.

THE SUB-MICROSCOPIC STRUCTURE OF "TICONAL" G MAGNET STEEL

620.18:669.15.255.24.71.018.58:621.318.2

"Ticonal" G is an alloy for permanent magnets, composed of 51% Fe, 24% Co, 14% Ni, 8% Al and 3% Cu [1]). When this alloy is cooled at a predetermined rate in a magnetic field and then annealed, an anisotropic material with excellent magnetic properties is obtained [2]): the value of $(BH)_{max}$ for this material amounts to about 5.0×10^6 gauss.oersted, measured in the direction of the field during cooling. Even better results can be reached with single crystals of this material (which crystallizes in the cubic structure) if it is ensured that during cooling down a (100)-direction is parallel to the direction of the field [3]). The $(BH)_{max}$ can thus attain a value of 8.0×10^6 gauss.oersted.

Fig. 1 is a photo-montage giving a three-dimensional impression of the metallographic structure of a single crystal of this kind. It has been obtained by placing electron-microscope photographs of the structure of the (100)-planes on the faces of a wooden cube, the direction of the magnetic field during cooling being indicated by the arrow. The photographs concerned are shown separately in *figs. 2* and *3*.

They reveal a substantial difference between the structures of a (100)-plane parallel to and perpendic-

[1]) The alloy known in British and American literature as Alnico V is an identical alloy.

[2]) B. Jonas and H. J. Meerkamp van Embden, Philips tech. Rev. **6**, 8-11, 1941.

[3]) Netherlands patent No. 71925; see Philips tech. Rev. **18**, 358-360, 1956/57 (No. 12).

ular to the magnetic field. The former is characterized by elongated elements with the longest dimension in the direction of the magnetic field (fig. 2). The (100)-planes perpendicular to the magnetic field during cooling are characterized by a structure in which mutually perpendicular elements predominate which, however, exhibit no pronounced elongated form (fig. 3).

Fig. 1. Photo-montage giving a three-dimensional impression of the metallographic structure of a single crystal of "Ticonal" G. The arrow indicates the direction of the magnetic field applied during the cooling period. This picture was obtained by placing the structural photographs (magnification 160 000 diameters) shown in figs. 2 and 3 on the corresponding faces of a wooden cube.

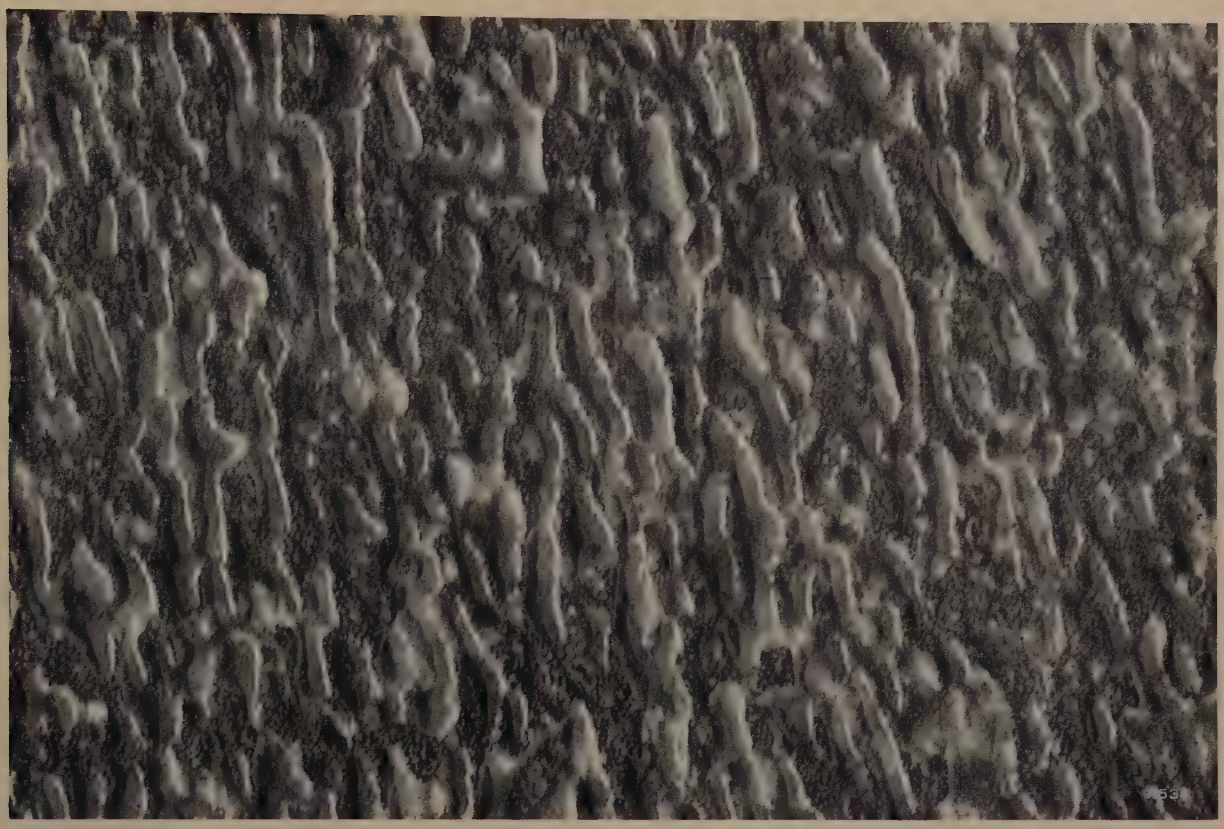


Fig. 2. Electron-microscope photograph of the structure of a (100)-plane *parallel* to the external magnetic field applied during cooling. Direct carbon replica of a platinum-shadowed, electrolytically polished surface, etched with a dilute solution of nitric acid in alcohol. Magnification 160 000×. This photograph was made with the Philips electron microscope EM 100 b.

Fig. 3. As fig. 2, but for a cube face *perpendicular* to the magnetic field.

In describing the structure, which indeed corresponds to the earlier formed concept based on the magnetic behaviour of this material, the term "structural elements" has been used deliberately, leaving open the question whether a well-defined second phase has been precipitated, or whether we are concerned with a pre-precipitate stage characterized merely by periodic fluctuations in the composition. The structural pictures obtained do not indicate which of these two ideas is correct, since even with a gradually changing composition sudden changes in the susceptibility to chemical attack may occur, such that etching produces discontinuities which resemble phase boundaries (Tammann's boundaries of chemical resistivity).

The extremely fine structural details between the relatively coarser "grains" have not yet been satisfactorily explained.

It is not possible within the space of this short report to compare the present results with those obtained by other workers [4]). Another paper, now in preparation, will go further into these matters.

H. B. HAANSTRA, J. J. de JONG and J. M. G. SMEETS.

[4]) R. D. Heidenreich and E. A. Nesbitt, J. appl. Phys. **23**, 352-365, 1952.
K. J. Kronenberg, Z. Metallkunde **45**, 440-447, 1954.
H. Fahlenbrach, Tech. Mitt. Krupp **12**, 177-184, 1954; **14**, 12-15, 1956 (No. 1).
D. Schulze, Exp. Techn. Physik **4**, 193-204, 1956 (No. 5).

THE JUNCTION TRANSISTOR AS A NETWORK ELEMENT AT LOW FREQUENCIES

I. CHARACTERISTICS AND *h* PARAMETERS

by J. P. BEIJERSBERGEN, M. BEUN and J. te WINKEL.

621.375.4

Transistors of various types are now becoming available in large numbers. In Europe, the Philips semiconductor factory at Nijmegen (to name one centre), which started up on the 12th July 1955, has been working to capacity for some time. Equipment embodying transistors has on several occasions been dealt with in this periodical. It seemed to the editors desirable to give an outline of the elementary properties of the transistor as a network element, in order to provide a basis for later articles on transistor applications. The article which follows, and which falls within the framework of the series of transistor articles started in 1955, is the first part of this outline.

Transistor characteristics

The physical basis of the functioning of the junction transistor at low frequencies has been discussed at length in two articles published earlier in this Review [1] [2]). We shall now turn to the characteristics of the transistor at low frequencies, leaving the physical mechanism of the device out of consideration. We look at the transistor as a network element having three connecting wires, marked by the manufacturer as the emitter, base and collector connections. Hence three different currents and three different voltages are involved, a total of six variables (*fig. 1*). If we apply direct voltages between collector and emitter (V_{ce}) and between base and emitter (V_{be}) (by so doing we fix the third voltage V_{cb} between collector and base, for, in accordance with Kirchhoff, $V_{ce} = V_{cb} + V_{be}$), the three currents will have values determined by the voltages [3]). As regards the three currents, we know from Kirchhoff's first law that $I_c + I_b + I_e = 0$. If we plot two of them, e.g. I_b and I_c, as functions of V_{ce}, for various fixed values of V_{be}, we shall have the properties of the transistor in the form of two families of (static) characteristic curves (*fig. 2*).

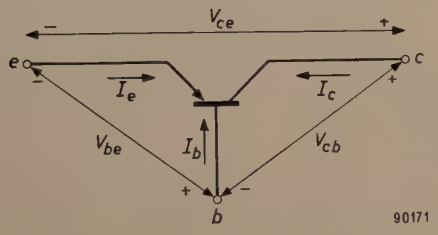


Fig. 1. The three voltages and three currents occurring as variables in transistors. *e* is the emitter, *c* the collector and *b* the base terminal. In accordance with the usual convention, the emitter terminal has an arrowhead to distinguish it from the collector terminal. The direction of the arrow shows whether the transistor is a *P-N-P* or *N-P-N* type; in the former case it points towards the base, in the latter case away from it.

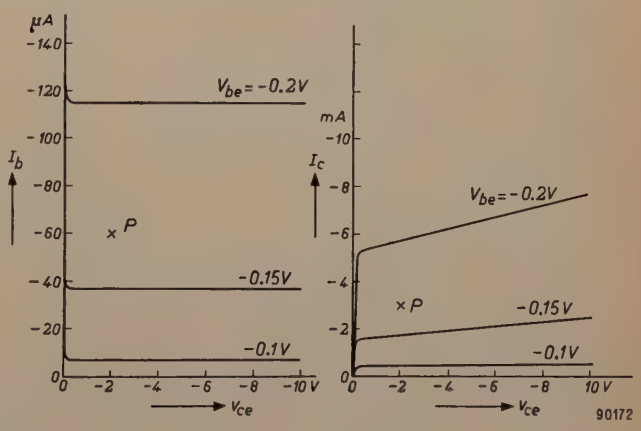


Fig. 2. Example of a pair of families of curves completely describing the properties of a transistor — here the Philips OC 71 — at low frequencies. The OC 71 being a transistor of the *P-N-P* type, all the quantities in the above graphs turn out to be negative. The biassing of the transistor might give an operating point at *P*, for example.

It is hardly necessary to point out that, when applied to A.C. signals, these static curves provide a valid approximation to the transistor properties only up to a certain limited frequency. Where that fre-

[1]) F. H. Stieltjes and L. J. Tummers, Simple theory of the junction transistor, Philips tech. Rev. **17**, 233-246, 1955/56.

[2]) F. H. Stieltjes and L. J. Tummers, Behaviour of the transistor at high current densities, Philips tech. Rev. **18**, 61-68, 1956/57 (No. 2).

[3]) The sequence of the indices (b for base, c for collector and e for emitter) implies a convention as to the sign of the voltages: for example, a positive value of V_{ce} signifies that the potential of the collector is higher than that of the emitter. Thus, for example, $V_{ce} = -V_{ec}$. Currents are regarded as positive when they flow towards the transistor — see the small arrows in fig. 1.

quency limit lies depends very largely on the type of transistor. For the OC 71, which we shall continually be taking as an example in this article, the limit may be placed at a frequency of 10 kc/s; for transistors that are specially designed for high frequencies the limit is very much higher.

The OC 71 is a transistor of the *P-N-P* type (see article [1])); accordingly, the transistor symbol in fig. 1 has the arrowhead that distinguishes the emitter from the collector pointing inwards. This is a convention in general use. Between *P-N-P* and *N-P-N* transistors there are no essential differences with regard to their behaviour as network elements.

Fig. 2 represents only two of the many possible families of curves. It is found, in fact, that 54 different families of curves can be drawn for one transistor [4]). Two mutually independent families of curves (i.e. such that one of the pair cannot be derived from the other) are sufficient to describe the transistor's properties at low frequencies. Which families of curves will be most appropriate and useful depends on circumstances. At this stage, suffice it to say that the pair shown in fig. 2 will seldom be encountered; the reason for their choice on this occasion is that the family of curves on the right corresponds to the usual anode characteristics of a tube, as will become clear in the next section.

In use, a transistor — like a tube — is biassed with certain direct currents and voltages. In any family of curves a point can be found that corresponds to these values — the operating point. It is of importance to observe that, when the transistor is biassed to a normal D.C. operating point (e.g. to *P* in fig. 2), the base current is always small in comparison with both the emitter current and the collector current (see the numerical values in fig. 2; for the physical reasons refer to article [1])). Hence the current that flows into the transistor at the emitter leaves at the collector almost unchanged in value. As to the voltages, in germanium transistors the voltage V_{be} between base and emitter is low in comparison with the voltages which those two electrodes have with respect to the collector. Thus $V_{bc} \approx V_{ec}$.

[4]) A. J. W. M. van Overbeek, Enige schakelingen met transistors, T. Ned. Radiogenootschap **19**, 231-260, 1954. We have $n = 6$ variables (the transistor voltages and currents) and require to know the number of ways in which any three of these may be selected and arranged between $m = 3$ locations (the abscissa, ordinate and parameter). This is given by $n!/(n-m)! = 120$ ways. Of these, the $3! = 6$ arrangements of the three voltages alone, and likewise the six arrangements of the three currents alone, merely give expression to Kirchhoff's laws and are therefore trivial, for they give no information about the transistor itself. There remain 108 graphs with meaningful content. However, this number is made up of pairs which are identical except that the ordinate of one graph is the abscissa of the other and *vice versa*.

A comparison with tubes

A triode, like a transistor, is a network element with three terminals; again, therefore, its properties can be expressed on paper in the form of two families of curves, the curves, for example, of grid current I_g and anode current I_a as functions of anode voltage V_{ak} (i.e. anode-cathode voltage), with the grid voltage V_{gk} (i.e. grid-cathode voltage) as the running parameter. The latter family constitutes the usual anode characteristics of a tube. A pentode has five terminals, and in theory there are many more current and voltage variables involved than in a triode or transistor. However, elementary treatment of the pentode is usually confined to the case in which the screen and suppressor grids have fixed potentials with respect to the cathode; in this way we arrive at the network element shown in *fig. 3*, and this again is one with three terminals. The properties of such a pentode can likewise be expressed in the form of two families of curves.

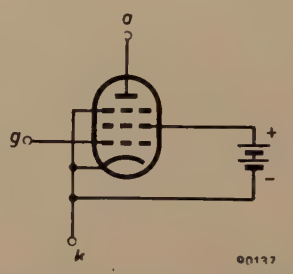


Fig. 3. A pentode has more than three terminals, but if screen and suppressor grids are given fixed potentials with respect to the cathode only three terminals are left, as is the case with a triode or transistor. Through the terminal marked *k* flows the sum of the cathode, screen-grid and suppressor-grid currents.

In certain respects transistors and tubes exhibit a considerable similarity in properties. The grid current of tubes is very small and hence the other two currents are roughly equal. Again, the voltage between grid and cathode is very low compared with the voltages between those electrodes and the anode, so that $V_{ga} \approx V_{ka}$. Thus, in comparing a transistor with a tube, the best analogy is obtained and the differences are brought out most clearly by regarding the base as corresponding to the grid, the emitter to the cathode, and the collector to the anode. Thus the curves plotted on the right of fig. 2 correspond to the anode characteristics of tubes.

One of the striking differences exhibited by tubes is that, in cases where the grid voltage is negative with respect to the cathode, the grid current is not merely small — it is quite negligible (the base current in a transistor, though small, is not negligible). Therefore, if we are only interested in applications in which the grid voltage remains negative (and this is so in receivers and amplifiers, but not usually in

transmitters), we have one variable less, and only one family of curves is required for expressing the properties of the tube. This is one of the reasons why it is often easier to understand the behaviour of tubes than that of transistors.

The common-emitter, common-base and common-collector configurations

We have seen that relations between only four of the six variables (two being voltages and two currents, as in fig. 2) suffice to describe a transistor completely. There are three simple cases where it is immediately clear which four variables it will be best to select; these cases arise when the transistor is inserted between two 2-terminal networks, in the manner shown in *fig. 4a*. Here the transistor itself is regarded as a 4-terminal network, or fourpole, that is to say a network in which two terminal pairs can be distinguished [5]). Like any other network element having three terminals, a transistor can be made into a fourpole in three different ways (figs. 4*b*, *c* and *d*). In fig. 4*b* the emitter has been given two terminals. One of them, together with the base, constitutes the first terminal pair; the other, together with the collector terminal, constitutes the second pair. Since the emitter connection is shared between the two pairs of terminals, the transistor is said to be in the common-emitter (also "*grounded* emitter") configuration. Alternatively, double terminals can be given to the base or to the collector, and doing so produces the common-base (fig. 4*c*) and common-collector (fig. 4*d*) configurations respectively.

In the common-emitter configuration, V_{be} and I_b are the voltage between and current through the one pair of terminals, and V_{ce} and I_c the voltage between and current through the other pair. Clearly, for this configuration we shall want to make use of those relations (i.e. those families of curves) that relate the voltages and currents just mentioned. It may easily be found from figs. 4*c* and 4*d* which currents and voltages should occur in the curves appropriate to the two other configurations.

The three simple circuits just described provide a suitable basis for discussing some of the fundamental properties of the transistor as a network element. In all three configurations a transistor can act as a

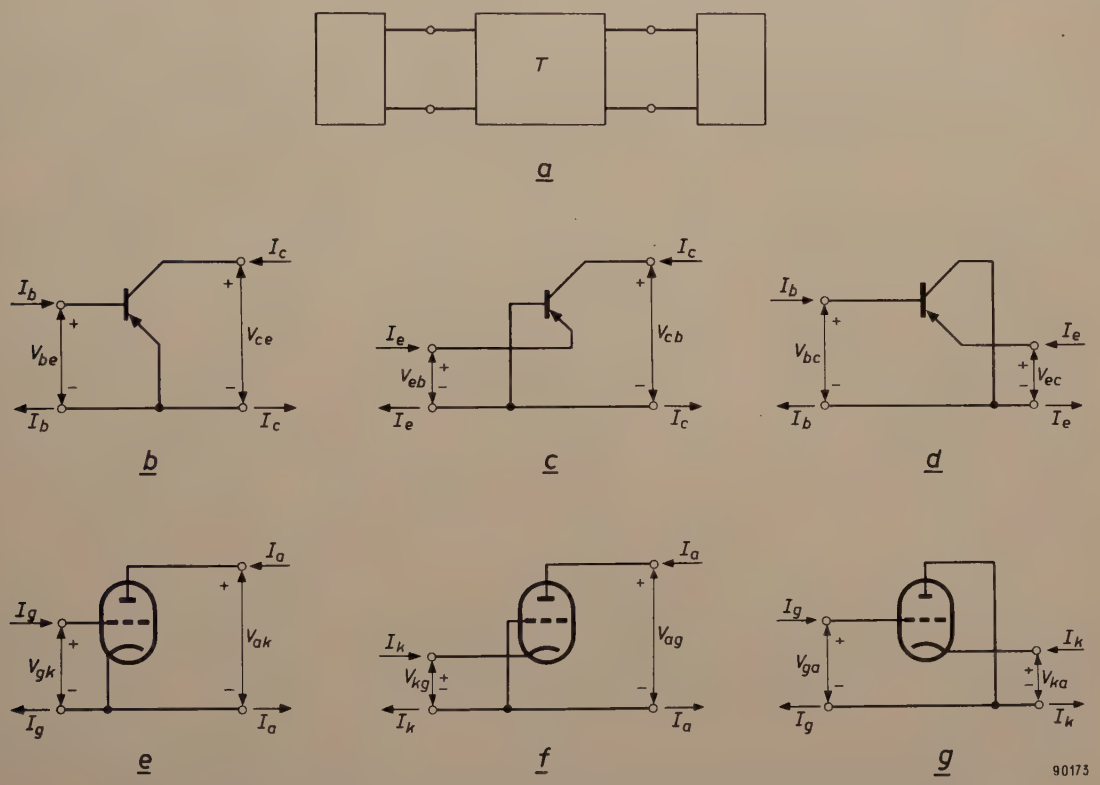


Fig. 4. If one of the three connections of a transistor T is provided with a double terminal, the transistor can be treated as a fourpole and inserted between two twopoles, as in (*a*). Three "configurations" are produced, (*b*) the common-emitter configuration (i.e. the emitter is common to input and output), (*c*) the common-base configuration and (*d*) the common-collector configuration, according to whether emitter, base or collector is given the double terminal. The corresponding tube circuits are also given — (*e*) common cathode (also "grounded" cathode), (*f*) common grid and (*g*) common anode (cathode follower).

[5]) Two terminals constitute a terminal pair when the current that enters by one of them leaves via the other.

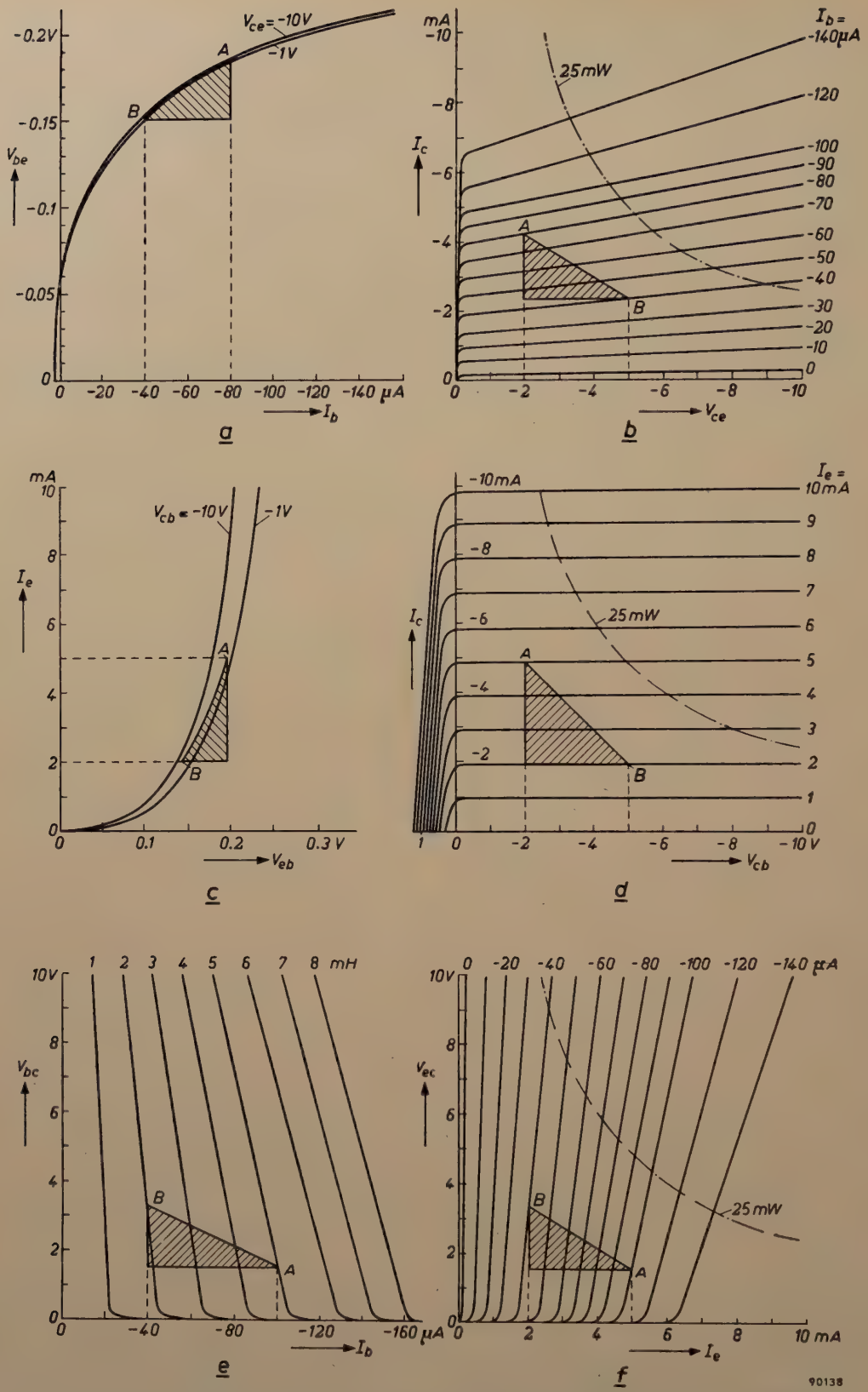


Fig. 5. Input characteristics (left) and output characteristics (right) of the OC 71 transistor in (*a*) and (*b*) the common-emitter, (*c*) and (*d*) the common-base, and (*e*) and (*f*) the common-collector configurations. These curves have been drawn to show that a power gain is obtainable in each of the three configurations. The load is imagined to be a pure resistance. *A* and *B* represent the points between which the instantaneous values of current and voltage vary as a result of the superposition of alternating current on the D.C. bias operating point. The areas of the hatched triangles on the left represent power taken up at the input; the areas of those on the right represent power delivered at the output. Differences in scale values should be noted. The hyperbolae in the output graphs are lines of constant power dissipation.

power amplifier. In the common-emitter configuration (fig. 4*b*), the direct voltage between and current through the base-emitter terminal pair are low in comparison with the direct voltage between and current through the other pair. This suggests that the same will be true of alternating currents and voltages superimposed on the D.C. operating point values. When base and emitter form the input pair and collector and emitter the output pair, therefore, we may expect power amplification to take place by way of simultaneous voltage and current amplification. This is confirmed by a glance at *figs.* 5*a* and *b*, representing the input and output characteristics for the configuration in question. If it is assumed that output voltage and current vary between the values represented by points *A* and *B* in fig. 5*b* and if the corresponding points *A* and *B* in fig. 5*a* are worked out, it is immediately clear that small variations at the input give rise to much larger variations at the output.

It will be of value here to put the corresponding tube circuit alongside the transistor configuration. In accordance with the preceding section, the corresponding tube circuit is that in which the cathode is the common electrode (this being the normal arrangement for a tube — see fig. 4*e* — viz. the common-cathode circuit or, more often, grounded-cathode circuit). In this the grid and cathode become the input terminals (small current and low voltage), and the anode and cathode become the output terminals (large current and high voltage).

In the common-base configuration (fig. 4*c*) the direct *currents* through the two terminal pairs (i.e. the emitter and collector currents) are of about the same magnitude; but the direct *voltage* between collector and base is much higher than that between emitter and base. That the same applies to alternating currents may be deduced from figs. 5*c* and *d*, by proceeding in the same way as above for figs. 5*a* and *b*. Thus, by choosing emitter and base as input pair and collector and base as output pair, power amplification is obtained as a consequence of voltage amplification. The tube circuit corresponding to this is the "grounded" grid (common-grid) circuit (fig. 4*f*). Here too, input and output currents have roughly the same magnitude, but the output voltage is much higher than the input voltage.

Finally we come to the common-collector configuration (fig. 4*d*). In this the voltages between the two pairs of terminals are roughly equal, but the currents differ. Power amplification consequent on *current* amplification is to be expected from this configuration. Base and collector are taken as the input terminals (the small base current becoming the input current), and emitter and collector as output terminals (the much larger emitter current becoming the output current). Figs. 5*e* and *f* confirm that a current amplification is obtained. The corresponding tube circuit is the common-anode circuit (better known as cathode-follower circuit — see fig. 4*g*), in which it is easy to see points of resemblance.

The terms "common-emitter", "common-base" and "common-collector" are frequently applied to circuits more complex than that of fig. 4*a*. Thus, the circuit shown in *fig. 6*, one that arises very frequently

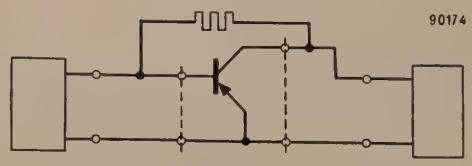


Fig. 6. A circuit frequently arising in practice, in which a (large) resistor is connected between collector and base. It is not possible here, without employing an artifice, to distinguish two pairs of terminals in the sense of footnote [5]).

in practice, is also referred to as a common-emitter arrangement; but it includes a resistor for negative feedback between collector and base, and it is not possible (at least, not without employing an artifice) to distinguish pairs of terminals in the sense of footnote [5]). In fact, the term "common emitter" is used in a wider sense to indicate that, in analyzing the circuit, use will be made of the currents and voltages associated with that configuration, and of the relations between those currents and voltages. The fourpole between two twopoles in fig. 6 naturally has characteristics different from those of the transistor standing alone. Changes in characteristics also result from modifications to the common-base and common-collector configurations and, *mutatis mutandis*, when modifications are made to the common-cathode, common-grid or common-anode circuits of tubes.

In what follows we shall mainly be concerned with the common-emitter configuration because this is the most commonly employed, albeit often in modified form (as in fig. 6, for example). In discussing this configuration we shall confine ourselves to variations of voltage and current so small that the transistor can be regarded as a linear network element.

The possible number of families of characteristics connecting the four interrelated currents and voltages in the common-emitter configuration is 12 (being $\frac{1}{2} \times 4!/(4-3)!$ — see footnote [4])). Most of them are used only incidentally, according to need and to taste, in order to investigate or explain the properties of transistors. It will be seen that the output characteristics in fig. 5*b* (where, in contrast to the usual practice with tubes, the input *current*, not the input *voltage*, is the

running parameter of the curve family) are remarkably straight and equidistant throughout a considerable range of currents and voltages. This signifies that within this range, provided that it is permissible to regard the input *current* as the input signal carrier (i.e. to regard the transistor as being "current driven"), the transistor introduces but little distortion. Against this, the input characteristics (fig. 5*a*) have considerable curvature; they are in fact exponential functions, as may be deduced from the physics of the transistor (see [1])). Hence, unless the amplitude of the input signal is very small, a sinusoidal input *current* implies a non-sinusoidal input *voltage*. Insofar as the input signal is an input *voltage*, therefore (i.e. when the transistor is "voltage driven"), considerable distortion may arise. For this reason it may be said that the transistor is first and foremost a current amplifier. In order to get an idea of the situation when the transistor is voltage driven — which in tubes is practically the only case arising — it will naturally be better to draw output characteristics with the input voltage as parameter, and these will give a less favourable impression of the linearity of a transistor. Even the characteristics of fig. 5*b*, however, do not remain equidistant at high values of I_b, but bunch together ever more closely. It is even possible to have non-linearity with respect to input *voltage* and non-linearity with respect to input *current* compensating each other to a considerable extent. The question of distortion in the cases where control of the transistor is neither by current alone nor by voltage alone, is a complex problem, the full treatment of which lies outside the scope of this article.

However, the reason why the input current is almost always taken as the running parameter in transistor output characteristics is connected not so much with distortion as with the way in which the transistor is biassed to the desired operating point. In choosing the operating point, use is made of the output graph (fig. 5*b*) because lines of constant power dissipation, i.e. hyperbolae of the form $I_c \times V_{ce}$ = constant, can easily be drawn in it. (The total power dissipated in the transistor actually amounts to $I_b \times V_{be} + I_c \times V_{ce}$, but the first term can be neglected.) In this way it can be seen immediately whether any operating point that may be chosen will involve excessive power dissipation. An impression is also gained of the extent to which current and voltage can be allowed to vary about the operating point without serious distortion arising. *Fig. 7* shows one possible method of biassing the transistor to the selected values of I_c and V_{ce}. The battery voltage V_0 (6 V, say) is high in comparison with V_{be}, the voltage between base and emitter (0.1 to 0.2 V), and hence $R_b \approx V_0/I_b$. To obtain a numerical value for R_b, one must be able to read off I_b from the graph; that is to say, I_b should be taken as the running parameter.

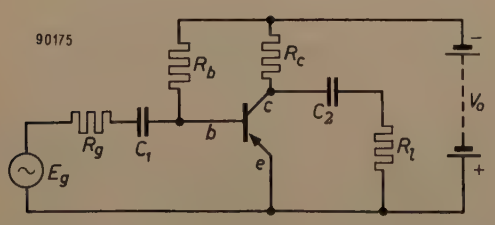


Fig. 7. A circuit with a transistor in the common-emitter configuration and with provision for the biassing. E_g is the e.m.f. and R_g the internal resistance of the signal source, R_l is the load, C_1 and C_2 are blocking capacitors, and R_b and R_c are resistors for setting the biassing. V_0: battery voltage.

Transistor fourpole parameters

If input voltage and current are denoted by V_1 and I_1, and output voltage and current by V_2 and I_2, the families of characteristics in figs 5*a* and *b* are graphical representations (for the transistor in the common-emitter configuration) of the relations:

$$V_1 = f(I_1, V_2), \quad \ldots\ldots \quad (1a)$$

$$I_2 = g(I_1, V_2). \quad \ldots\ldots \quad (1b)$$

It will become clear later on (page 22, column 1) why we have chosen I_1 and V_2 as the independent variables and V_1 and I_2 as the dependent variables. If current and voltage variations about the operating point are small enough for the relation between them to be regarded as linear, then we may write:

$$\Delta V_1 = \left(\frac{\partial V_1}{\partial I_1}\right)_{V_2} \Delta I_1 + \left(\frac{\partial V_1}{\partial V_2}\right)_{I_1} \Delta V_2,$$

$$\Delta I_2 = \left(\frac{\partial I_2}{\partial I_1}\right)_{V_2} \Delta I_1 + \left(\frac{\partial I_2}{\partial V_2}\right)_{I_1} \Delta V_2.$$

Introducing symbols whose definitions will be clear without further explanation, we can rewrite the above as follows:

$$v_1 = h_{11} i_1 + h_{12} v_2, \quad \ldots\ldots \quad (2a)$$

$$i_2 = h_{21} i_1 + h_{22} v_2. \quad \ldots\ldots \quad (2b)$$

The h parameters, as they are termed, occurring in these two equations, have meanings that can be deduced from the equations themselves. If v_2 is made zero, that is, if V_2 is kept constant (this amounts to short-circuiting the output for alternating current), it will be seen that:

$$h_{11} = \left(\frac{v_1}{i_1}\right)_{v_2=0} = \left\{ \text{Input resistance with output short-circuited.} \right.$$

$$h_{21} = \left(\frac{i_2}{i_1}\right)_{v_2=0} = \left\{ \text{Current amplification with output short-circuited; in these circumstances the usual term employed is "current amplification } \textit{factor}\text{".} \right.$$

If i_1 is made zero, that is, if I_1 is kept constant — and this amounts to open-circuiting the input for A.C. — we see that:

$$h_{12} = \left(\frac{v_1}{v_2}\right)_{i_1=0} = \left\{ \text{Feedback effect of output voltage on input voltage (reverse voltage amplification factor) with input open-circuited.} \right.$$

$$h_{22} = \left(\frac{i_2}{v_2}\right)_{i_1=0} = \left\{ \text{Output conductance (reciprocal of output resistance) with input open-circuited.} \right.$$

h parameters can of course be introduced for any fourpole, including each of the three possible configurations of a transistor. In what follows we shall, where necessary, indicate common-emitter h parameters by adding an index "e" to the h; thus, e.g., h^e. (Where the common-base or common-collector configurations are involved, we can write h^b and h^c [6].) The h parameters owe their name to the fact that a current and a voltage, I_1 and V_2, occur in them as independent variables, this giving them a certain *h*ybrid character.

Relations (2a) and (2b) remain valid when i_1, i_2, v_1 and v_2 are functions of time, providing they change so slowly — i.e. provided their frequencies are so low — that the relations (1a) and (1b) as given by the static characteristics continue to be obeyed at all times. In this article we shall only consider such low frequencies. It is quite possible to write relations of the form of (2) at higher frequencies, but then the h parameters are complex numbers with a complicated relationship to frequency. At low frequencies the h parameters are real numbers, as is apparent from their significance as the slopes of curves.

We started our reasoning with relations (1a) and (1b), taking I_1 and V_2 as independent variables, because we wanted to arrive at the h parameters. There are six different ways of choosing a pair of independent variables from amongst I_1, I_2, V_1 and V_2, and hence six sets of parameters are obtainable for the same fourpole. As always, it depends on the problem to be solved which choice will make calculation simplest. In the design of a simple amplifier equipped with transistors in the common-emitter configuration, the h^e parameters are the most convenient, as we shall see below. For this reason, and also because they can easily be determined by direct measurement with A.C. (this will likewise be gone into below — see page 25), the h^e parameters are often included in transistor data. It may be added that the various possible sets of parameters are not independent: given one set, one can calculate all the others [7]. h^e parameters for Philips OC 70 and OC 71 transistors at their normal biassing levels are given in *Table I*.

Table I. h^e parameters of Philips OC 70 and OC 71 transistors at normal biassing.

	OC 70 $V_{ce} = -2$ V $I_c = -0.5$ mA	OC 71 $V_{ce} = -2$ V $I_c = -3$ mA
h^e_{11}	2200 ohm	800 ohm
h^e_{12}	9×10^{-4}	5.4×10^{-4}
h^e_{21}	30	47
h^e_{22}	23×10^{-6} mho	80×10^{-6} mho

Characterization of a fourpole by less than four parameters

Various cases arise in which two parameters, or even one, suffice to characterize a fourpole, instead of the four that are generally necessary. This will be made clear by imagining a generator connected to an external resistance R_e, the generator supplying an e.m.f. of E_0 and having an internal resistance of R_0 (see fig. 8). The current delivered is $I = E_0/(R_0 + R_e)$. Since E_0/R_0 is the short-circuited current I_k, we can write:

$$\frac{I}{I_k} = \frac{1}{1 + R_e/R_0}. \quad \ldots\ldots \quad (3)$$

At the terminals of the generator, the voltage $V = IR_e$, so that

$$\frac{V}{E_0} = \frac{R_e/R_0}{1 + R_e/R_0}. \quad \ldots\ldots \quad (4)$$

In *fig. 8*, I/I_k and V/E_0 are plotted as functions of R_e/R_0, a logarithmic scale being employed for both ordinate and abscissa. When plotted in this way, one sees that it is possible to distinguish three regions of R_e/R_0, which are marked in fig. 8 as *I*, *II* and *III*. If R_e is small in comparison with R_0 (region *I*), the *current* is roughly constant and equal to the short-circuit current. If, on the other hand, R_e is large in comparison with R_0 (region *II*), it is the terminal

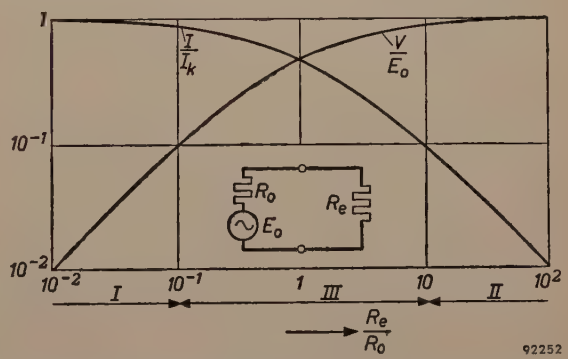


Fig. 8. Current and terminal voltage of a generator (e.m.f. E_0, internal resistance R_0) connected to an external resistance R_e, as functions of the ratio R_e/R_0. The scales of both axes are logarithmic. In region *I* the current approaches the short-circuit value I_k. In region *II* the terminal voltage approaches E_0, the e.m.f. of the generator. In region *III* matters are more complicated.

[6]) Frequently a prime (′) is used to show that the common-emitter configuration is intended, the h parameters without prime being made to relate to the common-base configuration. The drawback of this system is that it prevents the h parameters being used for any sort of fourpole.

[7]) Tables of transformation formulae may be found in: R. F. Shea, Principles of transistor circuits, Wiley & Sons, New York 1953, page 335.

voltage that is roughly constant and nearly equal to the e.m.f. In the middle region (*III*) matters are more complicated.

The four cases where only one parameter need be considered

We shall now apply these simple considerations to a fourpole. At its input the fourpole has a signal source (e.m.f. E_g, internal resistance R_g), and across its output is a resistive load R_l. At the *input side*, therefore, we have a generator (the signal source) connected to an external resistance, this being formed by the *input resistance* r_i of the fourpole with its load R_l (r_i is a function of R_l). At the *output side* we have a generator (the fourpole plus the signal source taken as a whole) connected to an external resistance R_l. The internal resistance of this generator is of course the *output resistance* r_u of the fourpole (r_u is a function of R_g).

It may happen that both the input/source resistance ratio r_i/R_g and the load/output resistance ratio R_l/r_u fall within region *I*. In this case i_1, the input current of the fourpole, is equal to the short-circuit current of the signal source and is thus, as far as the fourpole is concerned, a quantity determined by external conditions. i_2, the output current of the fourpole, is equal to the short-circuit output current, and from (2b) — v_2 now being approximately zero — is equal to $h_{21}i_1$. In this case, therefore, it is only necessary to know h_{21}, the current amplification factor, in order to describe the behaviour of the fourpole [8]). As we shall show by means of an example (pp. 23-24), this case arises when transistors in the common-emitter configuration are directly coupled in cascade. This explains why it is customary in transistor work to place so much emphasis on the current amplification factor; at the same time it is a good reason for including h_{21} among the set of fourpole parameters chosen to characterize the transistor. From the definition of the current amplification factor, viz. $h_{21} = (\partial I_2/\partial I_1)_{V_2}$, it can be seen that in order to arrive at such a set (including h_{21}, that is), I_1 and V_2 must be chosen as independent variables.

The three other cases in which one parameter suffices may be summarized as follows: input in region *II*, output in region *I* (fig. 8); both input and output in *II*; and finally, input in *I* and output in *II*.

Closer investigation of these cases shows up one of the essential differences between transistors, pentodes and triodes in sharp relief. In the case "input in *II*, output in *I*", the input voltage v_1 is a given quantity, being the e.m.f. of the signal source, while the output current is equal to the short-circuit value. Since this latter is given by $i_2 = v_1(\partial I_2/\partial V_1)_{V_2}$, the behaviour of the fourpole is wholly determined by the parameter $(\partial I_2/\partial V_1)_{V_2} = S$, the slope. This is e.g. the case arising with pentodes in the usual grounded-cathode circuits, and it is in fact usual with pentodes to place most emphasis on the slope of the tube. As may be seen from the definition of S, the set of fourpole parameters involving it is obtained by choosing V_1 and V_2 as independent variables.

In the case "both input and output in region *II*", v_1 is equal to the e.m.f. of the signal source, as it is in the previous case, and is again a given quantity. The output voltage has the value appropriate to the open-circuited output, when, in consequence, i_2 is approximately zero; that is, I_2 is constant. Here, therefore, $v_2 = v_1(\partial V_2/\partial V_1)_{I_2}$. The important parameter is here $(\partial V_2/\partial V_1)_{I_2} = \mu$, the voltage amplification factor. The case frequently arises in conventional triode circuits, where, of course, the internal resistance is much lower than with pentodes. It is clear that μ belongs to the set of parameters obtained when V_1 and I_2 are chosen as independent variables.

In the fourth and last straightforward case, that of "input in *I* and output in *II*", $v_2 = i_1(\partial V_2/\partial I_1)_{I_2}$ and hence $(\partial V_2/\partial I_1)_{I_2}$ is the important parameter; it belongs to the set obtained when I_1 and I_2 are chosen as independent variables. This case may arise in a grounded-grid triode circuit.

A case where only two parameters need be considered

Of the cases in which only *two* fourpole parameters are sufficient to describe the behaviour of the fourpole, we shall discuss only that in which the input/source resistance ratio is in region *III* and the load/output resistance ratio in region *I*. It often arises when transistors in the common-emitter configuration are employed in simple amplifiers with resistance-capacity coupling between stages (see page 25). The output is effectively short-circuited ($v_2 \approx 0$) and the output current is therefore $i_2 = h^e_{21}i_1$. The statement that the input/source resistance ratio falls within region *III* expresses the fact that the input resistance of the transistor ($= h^e_{11}$, the value for short-circuited output, since v_2 is approximately zero [9]), see eq. (2a)) is not negligible in comparison with the source resistance R_g. Knowing this, we must calculate the input current from $i_1 = E_g/(R_g + h^e_{11})$. In order therefore to know what we can expect from our transistor, we must know both h^e_{11} and h^e_{21}, these being respectively the input resistance with output short-circuited and the current amplification factor. Again the important parameters are *h* parameters.

Input and output resistance, current gain, voltage gain and power gain

Calculation and graphical representation

It will have become clear from the foregoing that r_i and r_u, the input and output resistances of a fourpole, have an important bearing on its behaviour.

[8]) Strictly speaking, this reasoning need not always be valid: its validity can be upset by the feedback effect expressed by h_{12}. We do not want to interrupt the argument at this point for the sake of this refinement, but shall return to it later on.

[9]) The reservation made in footnote [8]) must again be made here.

If the fourpole is specified by its h parameters (or by any other set of parameters), it is an easy matter to work out r_i and r_u. r_i is the equivalent resistance of the twopole shown in *fig. 9a*. From the figure we see that

$$v_2 = -i_2 R_l. \quad . \quad . \quad . \quad . \quad . \quad . \quad (5)$$

By eliminating i_2 and v_2 from this equation with the help of (2a) and (2b) we obtain r_i $(= v_1/i_1)$ as a

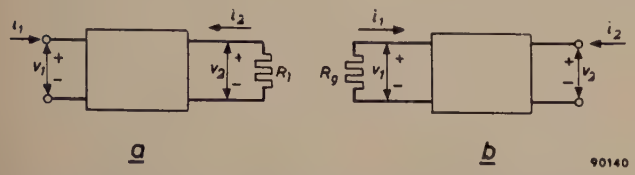


Fig. 9. *a*) Circuit for calculating the input resistance r_i of a fourpole having a resistance R_l as load. *b*) Circuit for calculating the output resistance r_u of the same fourpole when the generator has an internal resistance of R_g.

function of the load resistance R_l [10]). The resulting formula is given in *Table II*, and a graphical representation for a practical case (a Philips OC 71 transistor in the common-emitter configuration) is given in *fig. 10a*.

(2b) gives r_u $(= v_2/i_2)$ as a function of the source resistance R_g (see Table II and fig. 10*b*).

Since we have in mind the employment of the fourpole as an amplifier, we have a particular interest in the current, voltage and power gains, denoted by A_i, A_v and G respectively. A_i $(= i_2/i_1)$ is obtained by eliminating v_2 from (2b) and (5). A_v $(= v_2/v_1)$ is obtained by eliminating i_1 and i_2 from (2a), (2b) and (5). Conventions as to the sign of voltages and currents being duly observed, the power absorbed by the fourpole at the input side is $P_1 = i_1 v_1$, and the power delivered at the output side $P_2 = -i_2 v_2$. Thus the power gain is $G = -A_i A_v$. The formulae for A_i, A_v and G are given in Table II, and their graphical representations for the OC 71 in the common-emitter configuration appear in fig. 10*a* as functions of R_l.

The relative complexity of the formulae in Table II is chiefly due to the feedback from the output to the input, which is expressed by h_{12}. Formulae for the case $h_{12} = 0$, which also appear in the table, show the simplification that would result were this effect

Table II. Left column: some important fourpole quantities expressed in terms of the h parameters and the load and source resistances R_l and R_g. p is an abbreviation for $h_{12}h_{21}/h_{11}h_{22}$. The quantities thus given are, in order: input and output resistance, r_i and r_u; current, voltage and power gain, A_i, A_v and G; and finally G_{max}, the maximum value of the power gain. Right column: formulae for the same quantities in the case where there is no feedback from output to input (i.e. when h_{12}, the reverse voltage amplification factor, is zero).

$r_i = h_{11}\left(1 - p\dfrac{h_{22}R_l}{1+h_{22}R_l}\right)$	$r_i = h_{11}$
$r_u = \dfrac{1}{h_{22}\left(1 - p\dfrac{1}{1+R_g/h_{11}}\right)}$	$r_u = \dfrac{1}{h_{22}}$
$A_i = \dfrac{h_{21}}{1+h_{22}R_l}$	$A_i = \dfrac{h_{21}}{1+h_{22}R_l}$
$A_v = \dfrac{1}{h_{12}\left\{1 - \dfrac{1}{p}\left(1+\dfrac{1}{h_{22}R_l}\right)\right\}}$	$A_v = -\dfrac{h_{21}R_l}{h_{11}(1+h_{22}R_l)}$
$G = \dfrac{h_{21}}{h_{12}} \cdot \dfrac{ph_{22}R_l}{(1+h_{22}R_l)\left\{(1-p)h_{22}R_l + 1\right\}}$	$G = \dfrac{h_{21}^2}{h_{11}h_{22}} \cdot \dfrac{h_{22}R_l}{(1+h_{22}R_l)^2}$
$G_{max} = \dfrac{h_{21}^2}{h_{11}h_{22}} \cdot \dfrac{1}{(1+\sqrt{1-p})^2}$	$G_{max} = \dfrac{h_{21}^2}{4h_{11}h_{22}}$
(G_{max} occurs when $h_{22}R_l = 1/\sqrt{1-p}$)	(G_{max} occurs when $h_{22}R_l = 1$)

The output resistance r_u is the equivalent resistance of the twopole shown in fig. 9*b*. Here we may write:

$$v_1 = -i_1 R_g. \quad . \quad . \quad . \quad . \quad . \quad . \quad (6)$$

Elimination of i_1 and v_1 from equations (6), (2a) and

absent. In the second set of formulae input and output resistances are constants of the transistor, i.e. r_i is no longer dependent on the load, nor r_u on the signal source resistance.

Inferences to be drawn from the graphs

Let us imagine a directly-coupled cascade of three transistors, T_1, T_2 and T_3, of type OC 71, in the common-emitter configuration (*fig. 11a*), biassed to an operating point of $I_c = -3$ mA and

[10]) We shall confine ourselves to the cases where R_l and R_g are real (i.e. pure resistances). The h-parameters are also real at the low frequencies here assumed, and for this reason we shall also find that r_i and r_u have real values.

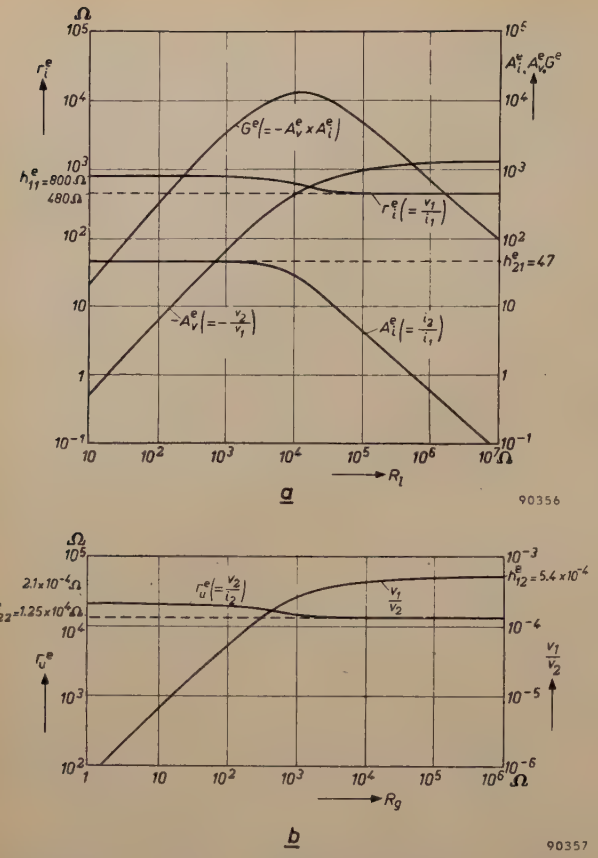


Fig. 10. Some important quantities plotted for the Philips OC 71 transistor in the common-emitter configuration (hence the e indices to the upper-right of symbols). All axes are on logarithmic scales. D.C. biassing (operating point): $I_c = -3$ mA, $V_{ce} = -2$ V.

a) Input resistance r_i^e, current gain A_i^e, voltage gain A_v^e ($-A_v^e$ is plotted here because, on account of the accepted convention, A_v^e is negative) and power gain G^e, as functions of the load resistance R_l.

b) Output resistance r_u^e and the reverse voltage amplification factor v_1/v_2, as functions of the source resistance R_g.

$V_{ce} = -2$ V. T_2 is therefore "sandwiched" between two other transistors. $r_u^e(T_1)$, the output resistance of T_1, is the source resistance $R_g(T_2)$ for T_2. Since $r_u^e(T_1)$ lies somewhere between 2.1×10^4 Ω and 1.25×10^4 Ω (see fig. 11*b*) — its exact value depends on the source resistance $R_g(T_1)$ — it follows from fig. 11*d* (hatched area) that $r_u^e(T_2) = 1.25 \times 10^4$ Ω. And since $r_u^e(T_2) = R_g(T_3)$, we find from fig. 11*f* that $r_u^e(T_3)$ has the same value of 1.25×10^4 Ω.

$r_i^e(T_3)$, the input resistance of T_3, lies somewhere between 800 Ω and 480 Ω (see fig. 11*g*), its exact value depending on its load, $R_l(T_3)$. Fig. 11*e* (hatched area) shows that, since $r_i^e(T_3) = R_l(T_2)$, then $r_i^e(T_2)$ is 800 Ω. In the same way, making use of $r_i^e(T_2) = R_l(T_1)$, we find from fig. 11*c* a value of 800 Ω for $r_i^e(T_1)$.

We learn from fig. 11*c* that the input resistance and current gain of T_1 have the values appropriate to short-circuiting of the output, in other words, they are equal to h_{11}^e and h_{21}^e respectively. Consequently these two h^e parameters are all we need to know about T_1. If data on the signal source are available, we shall be in a position to calculate the input current of T_1; this, multiplied by h_{21}^e (the current amplification factor), will give us the output current. The output current of T_1 is the input current of T_2. As we learn from fig. 11*e*, the current amplification of the second transistor T_2 is also h_{21}^e (value of $A_i^e(T_2)$ where curve is horizontal, see also fig. 10*a*); multiplying by this quite straightforwardly (here it is not necessary to know h_{11}^e) gives us the output current of T_2, which is, in turn, the

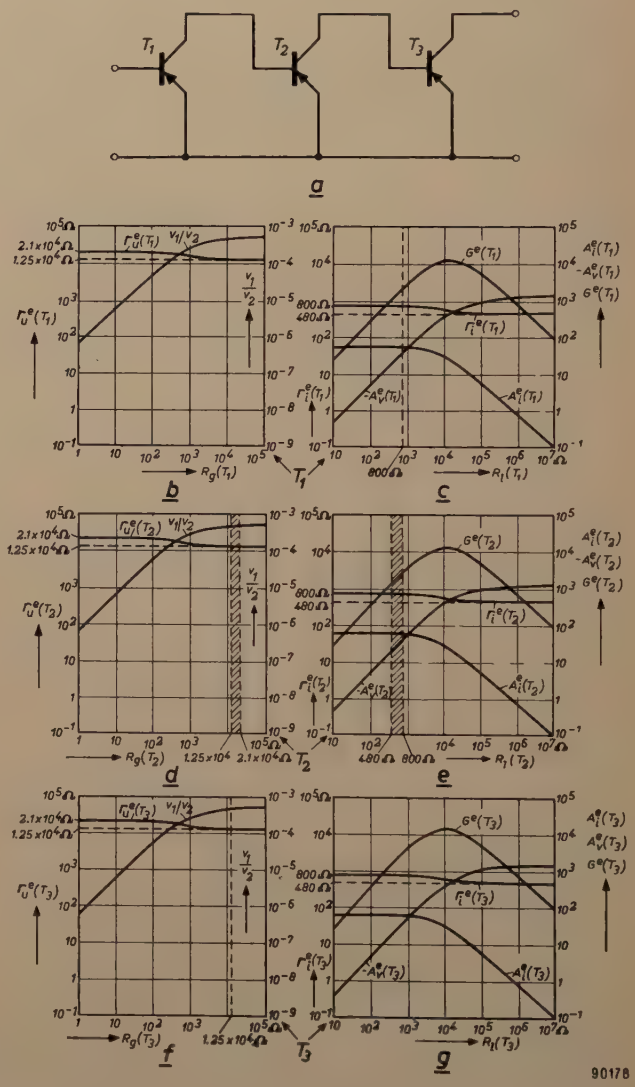


Fig. 11. Illustration of the situation obtaining when, as in (*a*), three transistors of type OC 71, T_1, T_2 and T_3, each in the common-emitter configuration, are directly coupled in cascade. The curves in (*b*), (*d*) and (*f*) are identical with those in fig. 10*b*, and relate to the outputs of T_1, T_2 and T_3 respectively. The curves in (*c*), (*e*) and (*g*) are identical with those in fig. 10*a*, and relate to the three inputs. The six graphs are annotated with such statements about the situation at the input or output in question as can be made without knowledge of the source resistance of T_1 or the load resistance of T_3.

input current of T_3. If, in addition, we know that $R_l(T_3)$ is lower than 1000 Ω, say, as in many cases it will be, fig. 11*g* shows that the current amplification will once again be equal to h^e_{21}. Multiplying by this once more, we obtain the output current of T_3. If, on the other hand, $R_l(T_3)$ is so large that A^e_i falls much below h^e_{21}, we should have to know h^e_{22} as well as h^e_{21} in order to calculate first A^e_i (using the formula given in Table II) and then the output current of T_3.

When resistance-capacity coupled amplifying stages are used (*fig. 12*), coupling resistors R_k occur across the outputs of the successive transistors. For simplicity we shall assume the capacitors to have so large a value that they offer no obstacle to alternating current. This means that the load resistance formed by the transistor of the following stage lies in parallel with the coupling resistor. The resultant load resistance is therefore diminished, and we can now say with still more certainty that the current gain and input resistance of the stage have the values appropriate to a short-circuited output, namely h^e_{21} and h^e_{11}. This being so, we can work out the output current of T_1 in the fashion explained above; but in the present case that current will not have the same value as the input current of T_2. In order to obtain the input current of T_2, we must calculate how the output current of T_1 is divided between the coupling resistor and the input resistance of T_2; for this it is necessary to know the latter, i.e. to know h^e_{11}. The same, of course, applies to the calculation of the input current of T_3, always assuming that $R_l(T_3)$ is less than 1000 Ω.

We could have reached the same conclusions by making use of fig. 8 which, with the aid of the numerical values found above, would allow us to investigate the input/source resistance ratio and the load/output resistance ratio of the three transistors.

The two examples just dealt with demonstrate the importance of both the current amplification factor, and the input resistance with the output short-circuited.

There is a further important point that can be clarified by the graphs of fig. 10: it is possible to measure the h^e parameters directly and without difficulty at low frequency, say 1000 c/s (at which frequency the static characteristics are still obeyed). In principle some of the measurements should be carried out in a circuit with open input and others with short-circuited output. An actual break in the input loop, like actual short-circuiting of the output, would, of course, make it impossible to apply the necessary D.C. biassing. However, it is necessary neither to make R_g particularly large nor R_l particularly small in order to have the effect of an open input and a short-circuited output, respectively. It will be seen from fig. 10 that for an OC 71 transistor it would suffice for R_g to be greater than 5000 Ω and for R_l to be less than 1000 Ω. Of course, when the parameters of an unknown transistor are being measured, these limits are also unknown in the first place. In such circumstances the measurements are repeated for increasing values of R_g and decreasing values of R_l, up to the point where v_1/v_2 and i_2/v_2 on the one hand and v_1/i_1 and i_2/i_1 on the other cease to change in value.

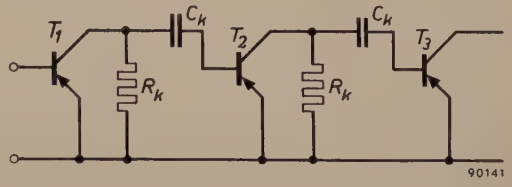


Fig. 12. Three transistors coupled by the resistors R_k and the capacitors C_k.

Internal feedback

In footnote [8]) a reservation was made concerning the conclusion that $i_2 = h_{21}i_1$ in the case where both the input/source resistance ratio and the load/output resistance ratio fell within region *I* (see fig. 8); in accordance with this conclusion — it was stated in the text — the current amplification factor was sufficient to describe the fourpole. The difficulty referred to in the footnote is the following. Suppose that R_g and R_l have values such that the input/source and the load/output resistance ratios do in fact both fall within region *I*. Now, if R_l (= R_e in fig. 8 when we consider the situation at the output) is reduced, the value of the load/output resistance ratio R_l/r_u (= R_e/R_0) is shifted further into region *I*, and consequently the value of i_2 is unaffected by the change. But r_i is a function of R_l. This is a consequence of the internal feedback, for if feedback is absent, i.e. if $h_{12} = 0$, it follows from the formula for r_i in Table II that $r_i = h_{11}$ = constant. If now r_i (= R_e in fig. 8 when we consider the situation at the input) increases as R_l decreases, then the value of the input/source resistance ratio r_i/R_g (= R_e/R_0) is shifted to the *right* in fig. 8. It is therefore conceivable that, as R_l approaches zero, the input/source resistance ratio moves out of region *I*. If so, i_1 ceases to be the short-circuit current of the signal source; it takes on a lower value. The value we would obtain for i_2 by multiplying the short-circuit current of the signal source by the current amplification factor h_{21} would therefore be too high. For transistors in the common-emitter configuration, r_i does indeed increase as R_l decreases (see the example of the OC 71 in fig. 10*a*), and one might justifiably think that the conclusion $i_2 = h_{21}i_1$ (for input/source and load/output resistance ratios both within region *I*) would be erroneous at least in some cases. If one goes deeper into the matter, however, it turns out that, at least for junction transistors, the conclusion is always valid for the case of both input/source and load/output resistance ratios lying within region *I*.

Similar considerations apply to the other cases dealt with on page 22. Strictly speaking, classification with the aid of fig. 8 of the various situations arising in fourpoles is only really possible if the feedback effect is neglected. With a vacuum tube, in the usual type of circuit (viz. common-cathode circuit), there is no feedback effect at low frequencies. This and

the fact that the input current in such a circuit is zero are the reasons why the treatment of tubes as linear, active, network elements is very much simpler than that of transistors.

Power gain and figures of merit of amplifiers

In general, an amplifier is employed whenever it is desired to have a power greater than is available without an amplifier. In most cases, however, the power gain G is by no means a good criterion for evaluating the performance of an amplifier in given conditions. The inadequacy of G becomes particularly clear in the extreme case of tube amplifiers (here we have in mind receiving tubes in the conventional common-cathode circuit). The input current is then negligibly small, and so in consequence is P_1, the power absorbed at the input. Since P_2, the power delivered, is finite, G is almost infinite. This certainly does not mean that any tube in any circuit is necessarily a good amplifier: whether the power delivered by the tube is greater than that which the signal source would have delivered without the intermediacy of the tube, remains an open question. And, in fact, reference is hardly ever made to power gain in connection with tubes. With transistors it is a different matter; since their input current is never zero, it is reasonable enough to talk about the power gain they give. Accordingly, in fig. 10*a*, G has been plotted as a function of R_l, this being a curve that may sometimes be useful in designing amplifiers. However, the G of transistors is still not a sure guide to the merits of the amplifier, for the question posed above remains unanswered. P_2 is given by GP_1; should P_1 be low because the transistor is badly matched to the signal source, it is quite possible that a higher value of P_2 might be made available by having better matching at the input, even though G should be lower.

A better figure of merit is arrived at by comparing the behaviour of an amplifier inserted between a signal source S and a load R_l, with that of an ideal matching transformer; in other words, by dividing the power that the amplifier delivers to R_l by the power that could be fed to R_l under the most favourable conditions, but without an amplifier. The figure thus obtained is called the "operating gain" or, alternatively, "transducer gain". We would invite special attention to the fact that *two* circuits are involved in the definition of operating gain, one circuit with an amplifier and one without. In fact, therefore, it is not a question here of a "gain of power" in the usual sense of the words — in the sense in which they are used in the term "power gain", for example. The sense is rather that of "advantage".

The concept of operating gain can reasonably be applied to any kind of amplifying device, including tubes, provided the amplifying device is considered in conjunction with the signal source and the load; it is not a quantity that can be stated for an amplifying device standing by itself. Indeed, the term itself says as much.

For a given signal source, the operating gain is at a maximum when amplifier and load are matched to each other. In the American literature, operating gain with matched output is called "available gain". It is possible to go a step further and also match the input: we then get the "maximum available gain". It is not difficult to see that maximum available gain is equal to G_{max}, the maximum power gain (see Table II), and is therefore a quantity that can be stated for an amplifying device standing alone, including a single transistor. For the OC 71 transistor in the common-emitter configuration it is approx. 10 000, i.e. approx. 40 dB.

To obtain the maximum available gain from a transistor is not impossible in practice, but doing so usually implies the use of matching transformers at input and output. The maximum available gain provided by tubes is almost infinite, but in view of the impossibility of matching the input, the fact is of no practical importance. In transistor circuits transformers will frequently be ruled out by considerations of space and price, the employment of one or more extra transistors being preferred.

Clearly, the comparison circuit need not necessarily be that in which the amplifying device is replaced by an (ideal) transformer. Sometimes the signal source directly connected to the load is chosen for this purpose. The figure of merit thus obtained is termed the "insertion gain". Here, as in "operating gain", the word "gain" has a sense different from the usual one.

The properties of a transistor as a linear active network element change with its biassing. This matter will be further investigated in a subsequent article, with the aid of an equivalent circuit whose elements are directly related to the physical functioning of the transistor (a "physical equivalent circuit"). In the same article the opportunity will be taken of discussing transistor equivalent circuits more generally. The temperature effects so important in transistor work will be dealt with in connection with the biassing in a third and final article.

Summary. The properties of a transistor at low frequencies can be expressed by pairs of families of curves, many different choices of these being possible. A transistor, like other network elements with three terminals (e.g. vacuum tubes), can be turn-

ed into a fourpole in three different ways, these being known as the common-emitter, common-base and common-collector configurations. Parallels with tube circuits are discussed. Where small signals only are involved, an active fourpole can be regarded as a linear active network element that can be characterized by four fourpole parameters. In simple cases the so-called h parameters are frequently selected for characterizing transistor behaviour. These parameters express the input voltage and output current in terms of input current and output voltage. The usefulness of h parameters is illustrated by examples, use being made of graphs in which input resistance and current gain are plotted as functions of load resistance, and output resistance is plotted as a function of source resistance. For transistors, unlike vacuum tubes in normal circuits, the concept of power gain is a usable one, but it is inadequate in itself as a criterion for the performance of an amplifier. A better guide is provided by "operating gain".

Two subsequent articles will deal with the effect of the biassing on the properties of transistors, and with temperature effects in connection with the biassing.

BOTTLING OF INJECTION FLUIDS

The Philips-Roxane Company markets various kinds of injection fluids [1]). The filling of the phials takes place under conditions of strict sterility, being carried out in a cupboard in which bacteria-free air is kept at a pressure slightly above atmospheric. Bacteria that may still find their way inside are killed by radiation from ultra-violet lamps fitted in the top of the cupboard. The phials are filled from a tube, which can be seen in the background. In the foreground the phials are being capped.

1) One of them being influenza-virus vaccine; see A. J. Klein and E. Hertzberger, Philips tech. Rev. **12**, 273-282, 1950/51.

PRE-MAGNETIZATION OF THE CORE OF A PULSE TRANSFORMER BY MEANS OF FERROXDURE

by H. G. BRUIJNING and A. RADEMAKERS.

621.314.2.073.3:621.318.124:621.374.3

Among the outstanding results yielded by recent solid-state research are two groups of non-metallic magnetic materials, ferroxcube and ferroxdure, which have both been the subject of several articles in this Review. Ferroxcube is a soft magnetic material with low losses even at high frequencies. Ferroxdure is a hard magnetic material, very suitable for permanent magnets. A combination of the two materials is found in new pulse transformers, developed for magnetron modulators. These transformers have a ferroxcube core, which is pre-magnetized by a ferroxdure magnet.

In electrical engineering and electronics we often find transformers whose primary or secondary current is entirely composed of pulses in the same direction. Some examples are: transformers feeding half-wave rectifiers, ignition coils for internal-combustion engines and the pulse transformers feeding radar transmitters. The core of such transformers is unfavourably loaded: the magnetic induction, instead of oscillating between $-B_{max}$ and $+B_{max}$, varies only between 0 and B_{max}, so that one half of the magnetization curve is not used at all. There is a means, however, to employ both halves, viz. by pre-magnetizing the core. We shall demonstrate in the following that this can be successfully effected by means of a *ferroxdure* permanent magnet [1]).

In this article we shall confine ourselves to the application of this principle in pulse transformers for radar transmitters.

Working principle of a radar transmitter

A radar transmitter radiates periodically wave trains of large power and short duration. The peak power may be of the order of hundreds of kW, the duration of each wave train 1 μsec, and the repetition frequency 1000 c/s. In order to operate the oscillator tube (as a rule a magnetron) intermittently, it is fed with an anode voltage of approximately square waveform.

Fig. 1 shows a circuit of the type commonly employed to generate such anode-voltage pulses. The thyratron *Th* being not ignited, the direct-voltage source E supplies a charging current flowing through the choke coil L, the pulse-shaping network P and the pulse transformer T. The pulse-shaping network P, which may be a lumped transmission line composed of coils and capacitors (delay line) is thereby charged. The magnetron M is connected to the secondary winding of T. By igniting the thyratron, the points A and C are suddenly short-circuited (short-circuiting of the D.C.-source is prevented by the choke L). This produces a square

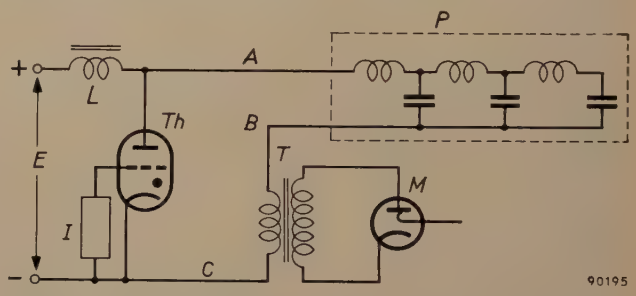


Fig. 1. Circuit for supplying the anode-voltage pulses for a radar transmitter. E direct-voltage source. L choke coil. *Th* thyratron. I ignition device for the thyratron. P pulse-shaping network (delay line). T pulse transformer. M magnetron.

voltage pulse across the primary of pulse transformer T, provided that the characteristic impedance of the delay line P is equal to the impedance between the points B and C. The duration of this pulse is equal to the time required by a wave front to travel through the network P and back again. During this pulse the energy stored up in P is transferred, via the pulse transformer, to the magnetron and partly radiated as high-frequency energy. At the end of each pulse the thyratron current falls to zero and the sequence can be repeated.

The pulse transformer

Frequency response

Magnetrons handling large powers require a high voltage (several ten-thousands of volts). This is produced by the pulse transformer T, which steps up a lower voltage supplied to the primary. A typical example: secondary 28 kV, 35 A; primary 7 kV, 140 A (peak values).

[1]) J. J. Went, G. W. Rathenau, E. W. Gorter and G. W. van Oosterhout, Philips tech. Rev. **13**, 194-208, 1951/52.

Fig. 2a shows the primary voltage v_1 as a function of the time t, and fig. 2*b* the desired form of the magnetic induction B in the transformer core during

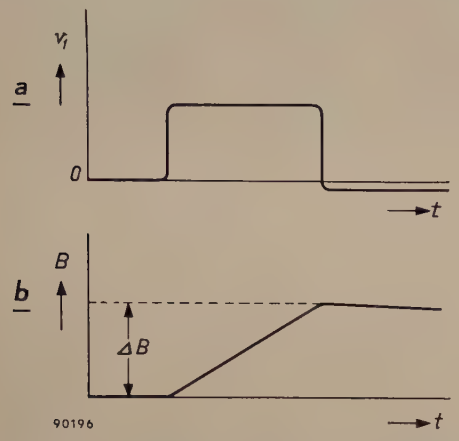


Fig. 2. *a*) Primary voltage v_1 of the pulse transformer T in fig. 1, versus time t. *b*) Magnetic induction B in the transformer core versus time.

the same period. The transformer must therefore transfer as faithfully as possible the entire frequency spectrum of v_1; this spectrum is shown (for relatively far too wide pulses, *fig. 3a*) in fig. 3*b*. The frequency-response curve of the transformer, however, will never be completely flat. Owing to the finite self-inductance of the primary, the induction B will not have the form of fig. 2*b*, but there will be a loss at the low frequency side, whilst the highest frequency that is adequately transformed will be determined by the required length of wire per turn. The frequency-response curve has, for example, a shape like that shown in fig. 3*c*, with the result that the secondary frequency spectrum will be of the form shown in fig. 3*d*. In the shape of the secondary voltage the loss of low-frequency components manifests itself as a "sagging" of the waveform (see fig. 3*e* and the oscillograms *fig. 4*) so that, in the wave train transmitted by the magnetron during a single pulse, the transmitted power diminishes accordingly. The loss of high-frequency components reduces the steepness of the leading and trailing edges of the pulse, so that the transmitted wave trains are less sharply defined. The first effect is more apparent if, for a given transformer core, too few windings have been used (self-inductance too low), and the second effect predominates if the total wire length and hence the number of turns is too large. A compromise is therefore to be sought.

Use of a ferroxcube core

The compromise attainable can be favourably influenced by suitable choice of the core material. A material of high permeability is advantageous in that it requires relatively fewer turns to give a specified self-inductance. A high saturation induction is favourable in view of the small diameter required for the core (small length per turn and hence shorter total wire length). If at the same time the eddy-current and hysteresis losses are low, the transformer will have a high efficiency.

High permeability, low losses and a not too low saturation induction are favourably combined in certain types of ferroxcube. From an engineering point of view, ferroxcube has the advantage over a metal core in that the former can be very easily composed from a few solid blocks; for a metal core only thin-rolled nickel-iron strip, wound into a ring, is suitable for this purpose. Ferroxcube moreover makes it possible to give different parts of the magnetic circuit different diameters, which is found to be of great advantage with pulse transformers; a ring

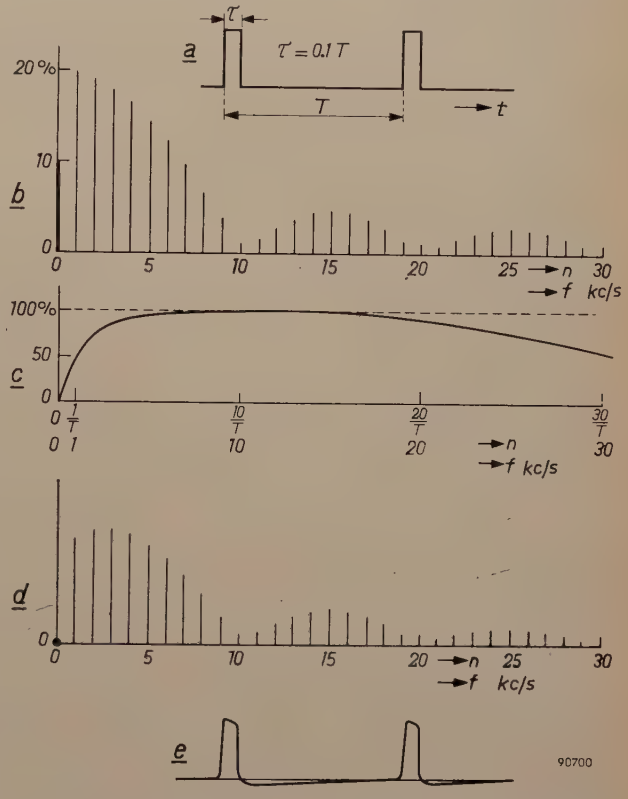
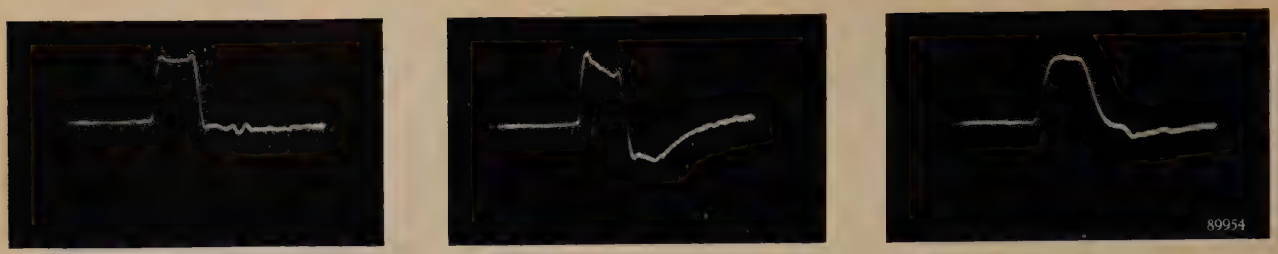


Fig. 3. *a*) Rectangular pulses of duration τ and repetition frequency $1/T$. With a view to the clarity of figs. (*b*) and (*d*), the pulses are shown relatively far too wide ($\tau/T = 0.1$ instead of approx. 0.001).
b) Spectrum of the pulse train (*a*). The amplitudes of the harmonics have been plotted vertically as a percentage of the pulse height; horizontally plotted are the order n and the frequency f of the harmonics, for the case $\tau = 10^{-4}$ sec, $T = 10^{-3}$ sec.
c) Frequency response of a pulse transformer.
d) Spectrum of the secondary voltage pulses.
e) The secondary pulses are distorted owing to the shape of curve (*c*).
(For the case of $\tau = 10^{-6}$ sec, $T = 10^{-3}$ sec, which is closer to normal radar usage, the scales of n and f should be multiplied by 100, and the line density in (*b*) and (*d*) should be 100× greater.)



a *b* *c*

Fig. 4. Oscillograms of the secondary voltage of a pulse transformer whose primary voltage consists of a series of nearly rectangular pulses (duration 1 μsec) and whose secondary winding is resistance-loaded.
a) Suitable shape of secondary pulses.
b) The too small primary self-induction causes a "sagging" pulse (loss of low-frequency harmonics).
c) Too long a wire length reduces the steepness of leading and trailing edges (loss of high-frequency harmonics).

wound from metal strip must necessarily retain a uniform diameter.

Ferroxcube 3C2, whose magnetization curve is shown in *fig. 5*, is particularly suited to this purpose. The saturation induction is about 0.5 Wb/m² (= 5000 gauss), corresponding to a magnetic field strength $H \approx 16$ kA/m (~200 oersted). If the material is to be employed at optimum efficiency, the variation ΔB should be enabled to occupy a long, steep section of the *B-H* curve. Suppose that ΔB be 0.3 Wb/m² and that this variation take place between the points *O* and *P* (fig. 5), then the relative permeability μ_r will be about 300. If the number of turns is so reduced that $\Delta B = 0.5$ Wb/m², and the curve *OQ* is chosen, then the increased yield of high frequencies (because of the smaller wire length) is offset by a loss of low frequencies, because μ_r drops to about 25, the self-inductance thus becoming considerably smaller.

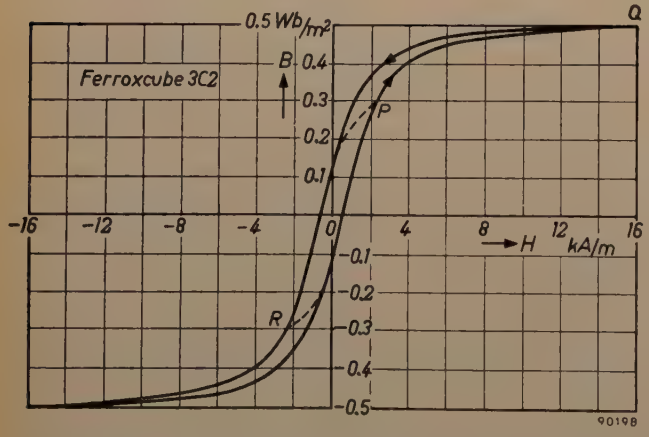


Fig. 5. *B-H* curve of ferroxcube 3C2. The curve *OP* corresponds to a large μ_r (= 300), but ΔB is only 0.3 Wb/m². If the curve *OQ* is chosen, then $\Delta B = 0.5$ Wb/m², but μ_r is reduced to 25. The condition with a pre-magnetization corresponding to −0.3 Wb/m² (point *R*) is a great deal more favourable: for the curve *RP*, $\Delta B = 0.6$ Wb/m² and $\mu_r = 300$.

Pre-magnetization

Conditions can be greatly improved if the steep part of the *B-H* curve below the *H*-axis is likewise employed. If, by applying a pre-magnetization of −0.3 Wb/m², *R* (fig. 5) is taken as the starting point instead of *O*, then the curve *RP*, corresponding to $\Delta B = 0.6$ Wb/m² and $\mu_r = 300$, is available. The application of pre-magnetization is therefore very profitable.

The question remains: in what way can pre-magnetization be most effectively established? The transformer core might be given a third winding, through which a constant direct current could be passed of such intensity that the core would be adequately pre-magnetized. This method has the advantage that it is possible to retain a *closed* magnetic circuit, so that the large permeability of the core material can be fully utilized. It has the drawback, however, of requiring a separate source of D.C.; moreover, the excitation circuit must have such a high impedance that it absorbs no appreciable amount of pulse energy.

A better solution is pre-magnetization by means of a *permanent magnet*, either of magnet steel or of ferroxdure. The latter has the advantage that it can pass an alternating flux without causing any appreciable eddy-current losses, owing to the high resistivity of this material.

Against the advantage that a permanent magnet does not require a D.C. source there is the drawback that the magnetic circuit must be interrupted for incorporating the magnet, thus rendering it more difficult to obtain a given self-induction. Particularly in the cause of ferroxdure, with its small reversible permeability, some difficulties in this respect would be expected. This drawback, however, is found to be a minor one: thanks to the high coercive force of

ferroxdure, a thin slab of this material is sufficient, so that the air-gap to be made in the transformer core can be narrow and consequently the decrease in self-induction remains moderate. We shall return to this later in this article. Let us first examine how the ferroxdure reacts to a current pulse through the transformer.

Behaviour of ferroxdure subjected to a pulse-shaped demagnetizing field

First of all we shall consider the general case of a pulse transformer with a core, pre-magnetized by a steel permanent magnet. The *B-H* curve of magnet steel is represented in *fig. 6*. The initial condition of the magnet is characterized by a point on the part of the *B-H* curve in the second quadrant, e.g. point *1*.

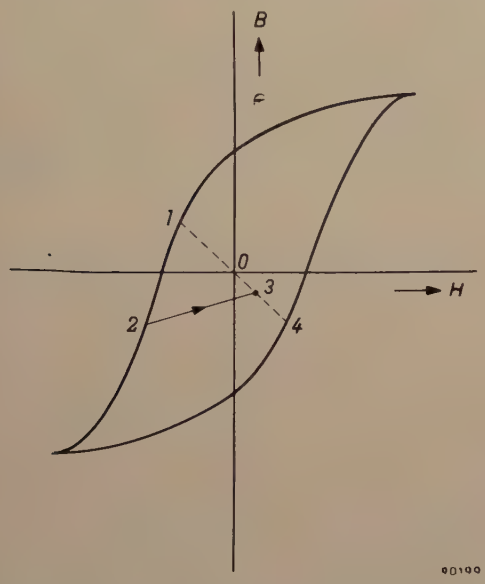


Fig. 6. *B-H* curve of a magnet steel. *1-0-4* working line of a magnet made of this steel, pre-magnetizing a pulse transformer. *1* initial working point. During a pulse the flux is reversed and point *2* is reached. At the end of the pulse, point *3* becomes the new working point. The pre-magnetization has thus taken on the wrong polarity.

If now the transformer is operated by a pulse of such amplitude and direction that it changes the sign of the induction, then the condition of the magnet will be temporarily represented by the point *2*, which lies equally far below the *H*-axis as *1* lies above it. The aim is obviously to have the initial condition restored after the pulse, but it will be readily appreciated that this is by no means certain. As soon as the pulse disappears, the working point traverses the line from *2* to *3*; the terminal point *3* being situated on the "working line" *1-0-4* of the system [2]). The line *2-3* has only a slight upwards slant (the "reversible permeability" of the magnetic material), so that *3* may be very well situated below the *H*-axis. This would mean that the pre-magnetization had not only changed its value, but also taken on the wrong sign, which, of course, has to be prevented at all costs.

With ferroxdure, however, matters are quite different. The reversible permeability of ferroxdure I is only little less than the slope of the *B-H* curve, so that the return curve lies closely along and nearly parallel with the *B-H* curve. The induction of the material, moreover, can be given a considerable negative value without the risk of losing the magnetization on the return path. An idealized representation of this behaviour, which we shall initially use in the reasoning which follows, is shown in *fig. 7*.

Approximated by straight lines, the magnetization J of ferroxdure I is plotted against H in fig. 7*a*. With increasing intensity of the demagnetizing field, J remains constant as long as H does not exceed the value ${}_J H_c$, but when H does exceed this value, J is entirely reversed. Fig. 7*b* shows the corresponding *B-H* curve: $B = J + \mu_0 H$, in which μ_0 is the permeability of free space ($= 4\pi \times 10^{-7}$ Wb/Am). When $\mu_0 H$ is plotted as abscissa on the same scale as B, as is done in fig. 7*b*, then the lines

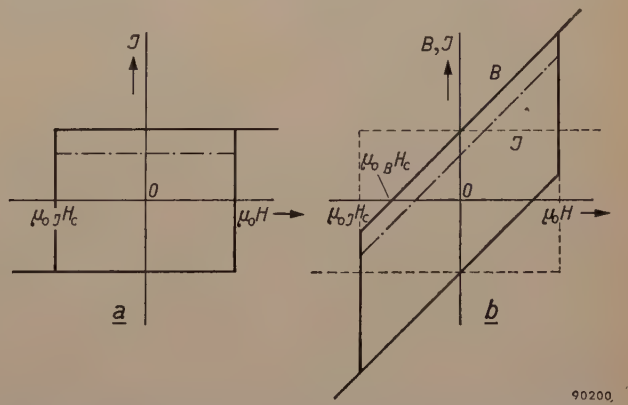


Fig. 7. *a*) Magnetization J, and (*b*) induction B of ferroxdure, as functions of $\mu_0 H$ (idealized). The value of J is shown as the dotted line in (*b*). The chain-dotted lines in both figures apply to a less strongly magnetized ferroxdure.

run at an angle of 45°. Only the upper slanting line is of interest to us here. It intersects the $\mu_0 H$-axis at a point $\mu_0 {}_B H_c$ (which, in ferroxdure, has a value significantly different from $\mu_0 {}_J H_c$). The working point can be shifted up and down along the upper sloping line of fig. 7*b* in a completely reversible manner. A condition of weaker magnetization is represented in both diagrams by a chain-dotted line, which may likewise be reversibly followed, and for which the same limiting value ${}_J H_c$ of the demagnetizing field applies.

[2]) Further in this article the working line will be more fully explained. See also A. Th. van Urk, Philips tech. Rev. 5, 29-35, 1940.

It should be observed that in this idealized picture the relative permeability $\mu_r = 1$, whereas the actual value for ferroxdure lies between 1.1 and 1.4. We shall deal with the actual B-H curve later in this article.

In our further study of the subject we shall assume that the designer of the pulse transformer has chosen a given primary self-induction L_1 and a given number of primary turns n_1; this also fixes the magnetic resistance (reluctance) R of the magnetic circuit, since if I_m is the magnetizing current and Φ_m the resulting alternating flux, the relations $I_m n_1 = \Phi_m R$ and $L_1 I_m = n_1 \Phi_m$ are valid, so that

$$R = \frac{n_1^2}{L_1}. \quad \ldots\ldots\ldots \quad (1)$$

R is the sum of the magnetic resistance R_{fxc} of the ferroxcube part of the circuit and of the magnetic resistance R_s of the air gap to be incorporated in the circuit to accomodate the ferroxcube magnet:

$$R = R_{fxc} + R_s .$$

For R_s we may put:

$$R_s = \frac{d}{\mu_r \mu_0 A_s},$$

d being the width of the gap (equal to the thickness of the ferroxdure slab) and A_s the area of the ferroxcube surfaces bordering the gap (A_s will as a rule be a good deal larger than the core diameter A_k of the ferroxcube within the coil). μ_r approximates to unity, both for ferroxdure and for air. Therefore, irrespective of the ferroxdure filling of the gap, we may put:

$$R_s = \frac{d}{\mu_0 A_s}. \quad \ldots\ldots\ldots \quad (2)$$

We shall assume here that not only R is fixed ($= n_1^2/L_1$, see (1)), but also the distribution of R into R_{fxc} and R_s, and thus also R_s itself. It follows from (2) that then also the ratio d/A_s ($= \mu_0 R_s$) is fixed.

We shall now attempt to determine the smallest admissible values of d and A_s, the smallest, that is, which involve no risk of permanent demagnetization. That this risk decreases with greater values of d and A_s (maintaining the same ratio), is attributable to two causes:

1) The larger the ferroxdure volume, the weaker the necessary pre-magnetization, thus allowing of a larger permissible negative B, according to fig. 7*b*.
2) The larger the area A_s, the smaller the variation of B resulting from a given flux variation.

These two factors fairly well define the critical condition where the risk of permanent demagnetization has to be seriously considered. It may be added here that, in practice, operation near this critical limit is not advisable. In case of certain faults the back flux may assume values considerably above normal, so that a certain safety margin is always necessary.

The faults in question are those whereby the magnetron fails to oscillate. The pulse transformer will than operate virtually without load and the earlier mentioned condition that the delay line (P, fig. 1) should be in series with a resistance equalling its own characteristic impedance is no longer fulfilled. As a result of this, the magnetizing current becomes many times (e.g. 40×) greater than normal. Fortunately, owing to the resulting saturation of the ferroxcube, the induction B_k in the core increases at a far lower rate than the current; however, the induction rises to the extent that a safety factor of 2-3 is indispensable to avoid demagnetization of the ferroxdure. This safety factor is defined as $(\Delta B_k)_{max}/2B_0$, $(\Delta B_k)_{max}$ being the maximum change in B_k that does not yet cause any demagnetization.

Let us first consider the case of an air-gap completely filled with ferroxdure (*fig. 8a*) and then that of a partly filled gap (fig. 8*b*).

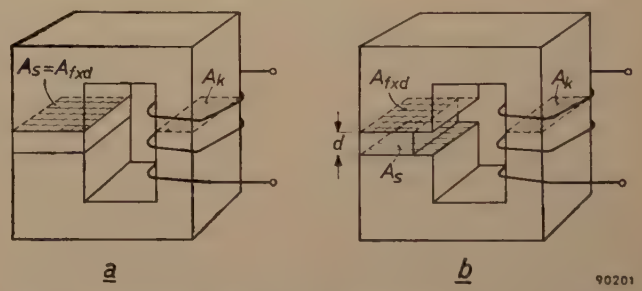


Fig. 8. Simplified shape of the ferroxcube core of a pulse transformer whose gap is filled completely (*a*), and only partly (*b*) with a pre-magnetizing permanent magnet of ferroxdure.

Completely filled gap

Since the pre-magnetizing flux Φ_0 may be assumed uniform throughout the circuit, we may write

$$H_{fxd} d = -\Phi_0 R_{fxc}, \quad \ldots\ldots \quad (3)$$

whilst

$$\Phi_0 = A_{fxd} B_{fxd},$$

the indices fxd and fxc standing for ferroxdure and ferroxcube respectively; for a completely filled gap $A_{fxd} = A_s$. From the two equations it follows that

$$\frac{B_{fxd}}{-H_{fxd}} = \frac{d/A_{fxd}}{R_{fxc}},$$

or, according to (2),

$$\frac{B_{fxd}}{-\mu_0 H_{fxd}} = \frac{R_s}{R_{fxc}}, \quad \ldots\ldots \quad (4)$$

which latter ratio, by assumption, is known.

In the idealized B-H curve of ferroxdure (*fig. 9*), equation (4) is represented by the straight line *0-1* (the working line). This line should be independent of the choice of d and A_s as long the ratio d/A_s is kept constant.

The intersection of the working line and the B-H curve defines the working point of the ferroxdure

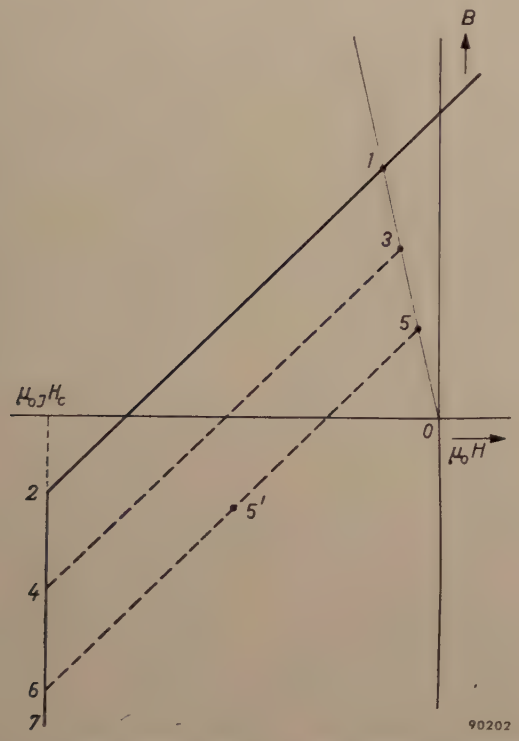


Fig. 9. *1-2-7* is part of the B-$\mu_0 H$ curve of strongly magnetized ferroxdure, *3-4-7* the same for a more weakly magnetized ferroxdure, and *5-6-7* for a still more weakly magnetized ferroxdure. *0-1* working line of a ferroxdure magnet incorporated in a magnetic circuit. Working point *1* is impracticable (safety factor less than *1*), *3* represents the critical case (safety factor = 1), and *5* permits a safety factor = 2.

when the transformer is currentless. The three parallel lines in fig. 9 are applicable to three different transformer sizes. Let us first choose d and A_s such that completely magnetized ferroxdure is necessary to produce the desired flux Φ_0 in the currentless transformer. This is the flux that brings out about an induction of 0.3 Wb/m² in the ferroxcube core and is represented by the point R in fig. 5. The working point would be situated at *1* (fig. 9). It is evident that if the flux is reversed, a point *2* would be passed, the latter being situated less below the horizontal axis than *1* lies above it. This case, therefore, cannot be used, and we have to increase A_s and d and use less strongly magnetized ferroxdure. If this is continued until the B value corresponding to point *3* is just adequate to produce the desired flux Φ_0, we arrive at the critical condition (safety factor 1): if the flux is reversed point *4* is reached, which lies equally far from the axis as *3*.

In practice, however, a certain safety margin is required, which is provided by choosing point *5* as working point. If the flux is reversed, *5'* becomes the working point (the ordinate of *5* is B_0, that of *5'* is $-B_0$). The reverse flux is now allowed to assume such a value that the point *6* with ordinate $-3B_0$ is reached, the safety factor being the ratio of the length *5-6* to the length *5-5'*, which is 2.

The procedure for determining the dimensions d and A_s is now a simple one. The pre-magnetizing flux needed to produce the desired pre-induction B_0 (= 0.3 Wb/m² in fig. 5) in the core is $\Phi_0 = A_k B_0$. Let B_w be the ordinate of the working point corresponding to a sufficiently large safety factor (such as point *5* in fig. 9), then the required cross-section A_{fxd} for the ferroxdure is given by:

$$A_{fxd} = \Phi_0 / B_w = A_k \, B_0 / B_w \, .$$

From this cross-section, by means of eq. (2), we find the necessary thickness d of the ferroxdure slab (this being also the width of the gap).

Partly filled gap

On the foregoing theoretical considerations we have arrived at a workable design for a transformer with a gap completely filled with incompletely magnetized ferroxdure (fig. 8*a*). The question now arises whether it is also possible (and perhaps better) to use a gap partly filled (fig. 8*b*) with more strongly magnetized ferroxdure. We find that the latter case can be easily related to the foregoing as follows.

If we reduce the cross-section of the filling in a given gap ($A_{fxd} < A_s$, fig. 8*b*), then obviously the magnetization of the ferroxdure must be raised to such an extent that the flux Φ_0 through the core remains the same. This means that the magnetic potential difference between the boundary planes of the gap, $\Phi_0 R_{fxc}$, must retain the same value. It follows from (3), therefore, that H_{fxd} must not change either. By analogy, we can also demonstrate that under conditions of reversed flux the value of H_{fxd} must remain the same. Let *a-b* in *fig. 10* be the earlier mentioned working line for a completely filled gap, then the working line for a gap appropriately filled with completely magnetized ferroxdure will be *c-d*, which lies parallel to and vertically above *a-b*. The corresponding change in induction ΔB is hence equally large in either case, since an equal flux variation $2\Phi_0$ is evenly distributed throughout the

gap, irrespective of what portion of the gap is taken up by air and what by ferroxdure (both being here assumed to have $\mu_r = 1$); in either case the change ΔB will be $2\Phi_0/A_s$.

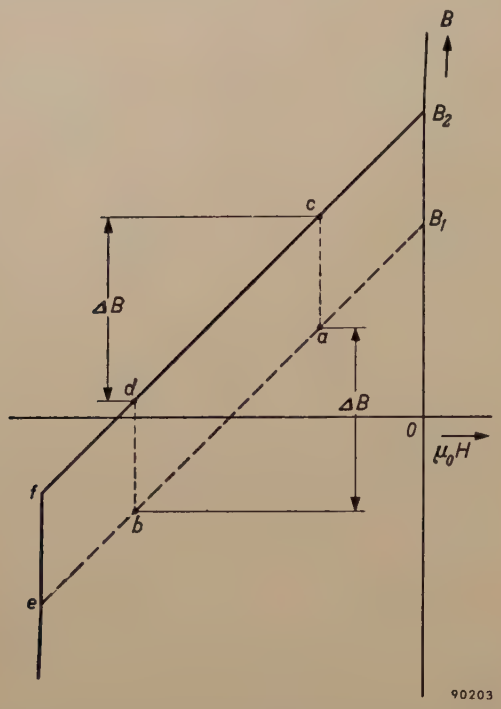


Fig. 10. *a-b* working line of a ferroxdure magnet completely filling the gap of the pulse transformer; remanence B_1. If the gap is partly filled with more strongly magnetized ferroxdure (remanence B_2), the working line becomes *c-d*, situated vertically above *a-b* and along the characteristic B_2-*f*, which runs parallel with B_1-*e*. The change in induction ΔB as well as the safety factor is the same in either case (at least in these idealized curves).

We further notice that the maximum range through which B can travel without causing permanent demagnetization likewise remains unaltered: accordingly the safety factor ($ae/ab = cf/cd$ in fig. 10) is also unaltered. The two cases depicted here are completely equivalent. Later we shall see, however, when the actual B-H curve is substituted for the idealized straight-line representation used up to now, that designs using the completely filled gap are to be preferred.

The question now arises to what extent the gap has to be filled with more strongly magnetized ferroxdure; in other words what must be the ratio $a = A_{fxd}/A_s$?

The equivalence of the two cases considered in fig. 10 is a result of the fact that the only relevant factor in this question (the ratio R_s/R_{fxc} being supposed constant) is the short-circuit flux of the ferroxdure, $A_{fxd}B_{rem}$, irrespective of the manner in which the gap is filled. Therefore, when a core with completely filled gap is modified into one with a gap partly filled with more strongly magnetized material, it is essential to keep $A_{fxd}B_{rem} = aA_sB_{rem}$ constant.

When changing over from *a-b* to *c-d* in fig. 10, we must choose $a = B_1/B_2$ (A_s not being changed), and for intermediate cases, where the remanence is B_{rem}, the working line will as a rule lie between the vertical dotted lines, provided that a and B_{rem} are so chosen that $aB_{rem} = B_1$.

This concludes our general theoretical treatment of the subject. We shall now consider the actual B-H curve and find out to what extent our conclusions have to be modified.

Actual B-H curve of ferroxdure

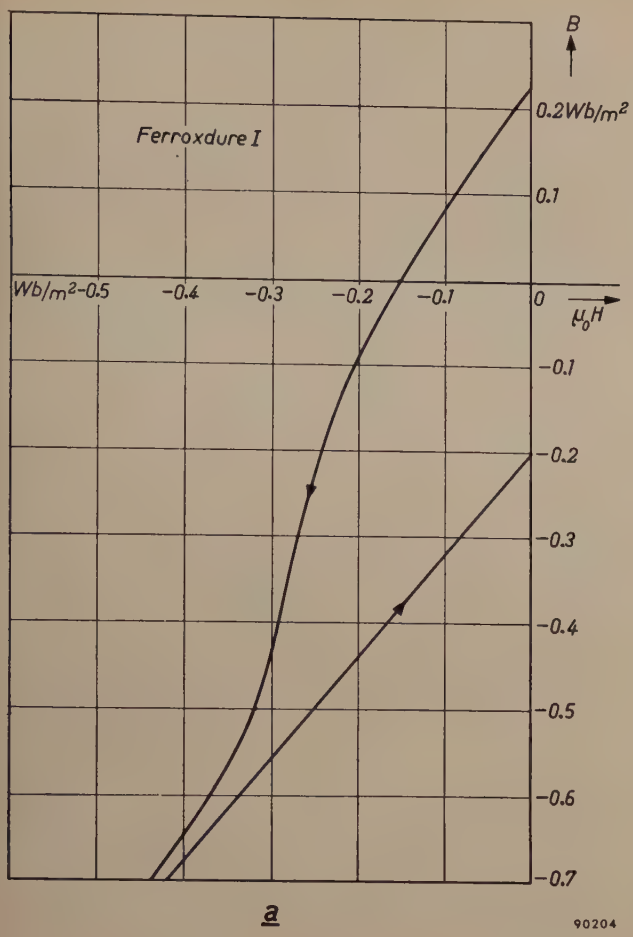
Fig. 11a shows the B-μ_0H curve measured on ferroxdure I. The part we are primarily concerned with is represented on an enlarged scale in fig. 11*b*, together with the reversible characteristics for $B_{rem} = 0.20$ and 0.12 Wb/m².

In the following we shall retain a constant ratio R_s/R_{fxc}, which we shall assume to be unity ($R_s = R_{fxc} = \frac{1}{2}R$), which was found to be a convenient value in practice. For a completely filled gap ($a = 1$), according to eq. (4), the working line in the B-μ_0H diagram will run at an angle of 45°.

Let us examine what possibilities these two conditions of magnetization may offer. With $B_{rem} = 0.20$ Wb/m², we find the working point *1*, where $B_{fxd} = 0.08$ Wb/m². In order to produce the desired pre-induction $B_0 = 0.30$ Wb/m² in the core, we must therefore choose $A_s = (0.30/0.08)A_k = 3.8\ A_k$. We find, however, from fig. 11*b* that the intersection *2* with the B-H curve lies less below the axis than *1* lies above it, so that reversing the polarity is not permissible. This is also manifest from the fact that the safety factor is less than 1, the value for $(\Delta B)_{max}$ in ferroxdure reading 0.13 Wb/m² in the diagram, so that the safety factor is as low as $0.13/(2\times0.08) \approx 0.8$.

The weaker pre-magnetization $B_{rem} = 0.12$ Wb/m² gives a far better result. The working line becomes *3-3'* in fig. 11*b*, with the value $B_{fxd} = 0.05$ Wb/m² in point *3*, so that A_s must become $(0.30/0.05)A_k = 6\ A_k$. The required quantity of ferroxdure, therefore, is $(6/3.8)^2 = 2.5$ times that in the previous example (both the cross-section A_s and the thickness d increase by a factor of 6/3.8, their ratio being constant according to (2)). This larger amount of material results in a remarkable improvement of the safety factor. This factor, being the ratio of the length of *3-4* to that of *3-3'*, now becomes 2.5, which may be considered adequate.

We now have a workable design with weakly magnetized ferroxdure filling the entire air-gap.



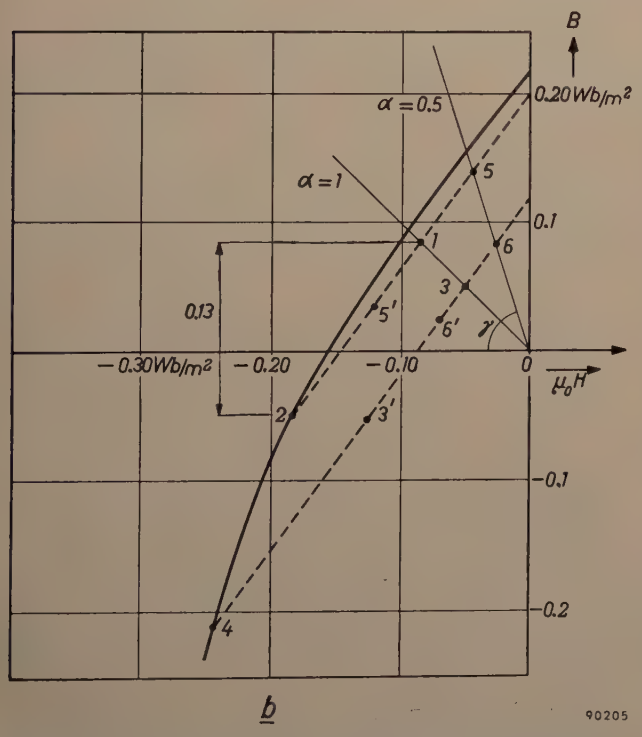


Fig. 11. *a*) B-$\mu_0 H$ curve measured on ferroxdure I. *b*) Part of (*a*) enlarged, with the reversible characteristics for $B_{rem} = 0.12$ and 0.20 Wb/m², and the working lines for $\alpha = 1$ and 0.5. Working point *1* cannot be used (safety factor less than 1). The curve *3-3'* corresponds to working point *3*, *5-5'* corresponds to *5*, and *6-6'* to *6*.

From this, we might now change over to a partial filling with stronger magnetic material, in the manner described earlier by shifting the working-line *3-3'* vertically upward, whilst appropriately varying the value of α.

We may, however, also approach the matter differently again, viz. by assigning a fixed value (< 1) to α and then investigating the possibilities thus offered. We shall do so for $\alpha = 0.5$ in combination with the values $B_{rem} = 0.12$ and 0.20 Wb/m².

Let us start by finding the working line for the currentless condition. For a half-filled gap we must replace R_s in the numerator of (4) by the magnetic resistance of the half gap, viz. $2R_s$ ($= 2R_{fxc}$); the denominator contains the magnetic equivalent resistance of the ferroxcube core shunted by the empty half of the gap: R_{fxc} in parallel with $2R_s = R_{fxc}$ in parallel with $2R_{fxc} = \frac{2}{3} R_{fxc}$. The slope of the working line for $\alpha = 0.5$ (see fig. 11*b*) is therefore $\tan \gamma = 2R_{fxc}/(\frac{2}{3}R_{fxc}) = 3$. This working line intersects the two reversible characteristics in the points *5* and *6*. In order to evaluate the induction B_0 in the ferroxcube core, the flux through the empty half of the gap (induction B_l) must be subtracted from the flux produced by the ferroxdure. In the empty half H equals H_{fxd}. We may therefore write:

$$A_k B_0 = \tfrac{1}{2} A_s (B_{fxd} + B_l) = \tfrac{1}{2} A_s (B_{fxd} + \mu_0 H_{fxd}), \quad (5)$$

in which B_l and H_{fxd} are negative and have values that can be read from fig. 11*b*. This permits us to evaluate the ratio A_s/A_k necessary to produce the pre-induction $B_0 = 0.30$ Wb/m². Thus we find the values $A_s/A_k = 6.7$ for point *5* and $A_s/A_k = 12$ for point *6*. The working points for the reverse flux — which we must know for determining the safety factor — can be found by substituting the above values of A_s/A_k and B_0 in (5). For the two cases considered here we arrive at the conditions

$$B + \mu_0 H = -0.09 \text{ and } -0.05 \text{ Wb/m}^2 \text{ respectively,}$$

corresponding to the points *5'* and *6'* in fig. 11*b*. The safety factor can now be read from the diagram: it amounts to (*5-2*)/(*5-5'*) = 1.8 for the material with $B_{rem} = 0.20$ Wb/m², and to (*6-4*)/(*6-6'*) = 5.0 for that with $B_{rem} = 0.12$ Wb/m².

In *Table I*, finally, we can summarize the cases dealt with in this section. For each of the four cases considered, are given the values of A_s/A_k, the short-circuit flux $\alpha B_{rem} A_s$ of the ferroxdure divided by A_k, the required volume of ferroxdure (as a percentage), and the safety factor.

Table I. Characteristic quantities relating to a pulse transformer core, pre-magnetized with ferroxdure I, for values of $B_{rem} = 0.12$ and 0.20 Wb/m² and for a completely-filled gap ($\alpha = 1$) and for a half-filled gap ($\alpha = 0.5$), R_s being equal to R_{fxc}.

	$\alpha = 1$		$\alpha = 0.5$	
B_{rem} (Wb/m²)	0.12	0.20	0.12	0.20
A_s/A_k	6.0	3.8	12	6.7
$\alpha B_{rem} A_s/A_k$ (Wb/m²)	0.72	0.74	0.72	0.67
Volume of ferroxdure (%)	260	100	530	160
Safety factor	2.5	0.8	5.0	1.8

Row 1 of the table shows that A_s/A_k is always fairly large, so that the ferroxcube surfaces in the air-gap must be substantially larger than the core diameter within the coil. This leads, as we shall see presently, to constructions different from those shown in fig. 8.

On page 34, on the strength of the idealized characteristic and on the assumption that $\mu_r = 1$ for ferroxdure, we have said that the short-circuit flux $\alpha A_s B_{rem}$ of the ferroxdure must be held constant. The actual values of this flux (divided by the constant cross-section A_k) as given in row 2 of the table, are indeed found to be fairly constant for the actual characteristic as well.

Rows 3 and 4 show the surprising result that the safety factor in the region considered is roughly proportional to the ferroxdure volume, and independent of its shape and state of magnetization (α and B_{rem}). It will, therefore, always be preferable to use a completely filled gap, since the ferroxcube yoke on either side of the gap can then have the smallest diameter, whereas on the other hand the saving in ferroxdure by using a partially filled gap is not worth mentioning. Of the four cases shown in the table that of column 1 is preferable for this reason. If a larger safety factor, e.g. 5, is required, it should be possible with $\alpha = 1$ to arrive at a better design than that of column 3. Owing to the inadequate safety factor, the case of column 2 cannot be used, whilst that of column 4 has a somewhat low safety factor and is, moreover, in view of $\alpha < 1$, by no means the best solution.

Example of an actual design

Finally, as an example we shall describe a pulse transformer operating in a laboratory set-up for testing magnetrons. The primary demands for this transformer were that it should produce a secondary pulse of the appropriate shape and that it should retain its pre-magnetization under all conditions.

Fig. 12 is a photograph of the core and the essential data are given in the caption. The completely filled gap is 2 mm wide and has an area $A_s = 8 \times 10 = 80$ cm², i.e. 8 × larger than the cross-section of the cylindrical cores. This large ratio A_s/A_k, together with the wish to use ferroxcube blocks of standardized dimensions, have led to the shape depicted here.

89953

Fig. 12. Core of the pulse transformer of a testing installation for magnetrons. Principal dimensions: base 100×170 mm, height 162 mm. Further principle data:

	Primary	Secondary
Peak voltage	7 kV	28 kV
Peak current	140 A	35 A
Impedance	50 Ω	800 Ω
Number of turns	16	66
Pulse duration	1 µsec	
Rise time	0.14 µsec	
Cross-section A_k within coil	10 cm²	
Area A_s of gap	80 cm²	
Width d of gap	2 mm	
Safety factor	approx. 3.	

The two vertical columns consist of ferroxcube 3C2 discs of 10 mm thickness (thicker discs are difficult to manufacture without cracks) and 36 mm diameter. The yokes are formed by rectangular blocks of ferroxcube 4B2, which can easily be manufactured in large pieces; if necessary, the somewhat smaller permeability can be compensated by giving the yokes a larger cross-section. Pre-magnetization is effected by eight 2×20×50 mm slabs of ferroxdure I, magnetized in the direction of their thickness. These are interposed between two large slabs

of ferroxcube, the upper one serving to concentrate the flux into the much smaller cross-section of the cylindrical part of the core. The safety factor is about 3. Each column is wound with a primary and a secondary winding, the primary windings being connected in parallel and the secondary windings in series. The whole assembly, clamped together by rods of insulating material, is immersed in oil. It should be noted that for a laboratory installation the dimensions are of secondary importance, so that no attempt has been made to minimize the dimensions of the transformer.

Summary. With transformers in which one of the windings carries unidirectional pulse-shaped currents, pre-magnetization of the core forms a means to arrive at a more favourable loading of the core. As an example the pulse transformer of a radar transmitter is discussed. To give the secondary pulse the appropriate shape, the transformer must have a considerable self-inductance, whilst the wire length of the turns must be kept short. Pre-magnetization of the core allows a more favourable compromise between these two conflicting demands. This can be effectively realized by incorporating in the magnetic circuit composed of ferroxcube blocks, a slab of magnetized ferroxdure. If it is arranged that the pre-magnetization brings about an induction of -0.3 Wb/m^2 in the ferroxcube within the coil, and that when pulses are applied to the transformer, this induction can fluctuate between -0.3 and $+0.3$ Wb/m^2, then the relative permeability of the core can reach a value as high as about 300.

Dimensions should be judiciously chosen to preclude the possibility of permanent demagnetization of the ferroxdure by the pulses. The theory of magnetic circuits incorporating a permanent ferroxdure magnet and subjected to a reversing flux is treated with the aid of an idealized straight-line magnetization curve, and subsequently verified with the aid of the actual curve. Finally a practical example of such a transformer is described.

ABSTRACTS OF RECENT SCIENTIFIC PUBLICATIONS BY THE STAFF OF N.V. PHILIPS' GLOEILAMPENFABRIEKEN

Reprints of these papers not marked with an asterisk * can be obtained free of charge upon application to the Philips Research Laboratory, Eindhoven, Netherlands.

2415: H. O. Huisman, A. Smit, P. H. van Leeuwen and J. H. van Rij: Investigations in the vitamin A series, III. Rearrangement of the retro-system to the normal system of conjugated double bonds in the vitamin A series (Rec. Trav. chim. Pays-Bas **75**, 977-1006, 1956, No. 7).

A new synthesis of vitamin A essentially based on the almost quantitative rearrangement of the retro-system to the normal system of conjugated double bonds is described. This rearrangement is carried out by converting the retro-C_{15}-acid and the retro-C_{20}-acid, by means of phosphorus trichloride in suitable solvents, into the corresponding retro-acid chlorides, which thereupon rearrange smoothly to the acid chlorides, with the normal system of conjugated double bonds, β-ionylidene acetic acid chloride and vitamin A acid chloride respectively. The acid chlorides are reduced with $LiAlH_4$ to the corresponding alcohols β-ionylidene ethanol and vitamin A. The physical and chemical properties of the intermediates and geometric isomers obtained in the various reaction steps are described.

2416: F. A. Kröger, G. Diemer and H. A. Klasens: Nature of an ohmic metal-semiconductor contact (Phys. Rev. **103**, 279, 1956, No. 2).

Models for an ohmic contact between a metal and a high-ohmic *n*-type semiconductor like CdS as proposed by Smith and Butler are not quite satisfactory with regard to experimental evidence. An alternative model is proposed, according to which a thin layer of the semiconductor adjacent to the electrode is strongly *n*-type, e.g. due to diffusion of trivalent metal ions from the electrode into the CdS or to the bombardment of the CdS surface, which is often used to clean the surface before applying the electrodes.

2417: P. A. Neeteson: The vacuum tube as a network component in pulse circuits (T. Ned. Radiogenootschap **21**, 171-185, 1956, No. 4).

In this article a survey and some illustrating examples are given of methods by which large-signal behaviour of vacuum tubes can be determined. Contrary to small-signal application, where the tube is operated by relatively small deviations around a fixed operating point in the conducting region, in large-signal operation the tube is rapidly brought from the cut-off into the conducting region and *vice versa*. This is essentially a switching action. Therefore, it is necessary to give some preliminary considerations on switch operation in networks. The method has proved to be useful in analyzing electronic pulse circuits. See also the book by the same author, announced in Philips tech. Rev. **18**, 284, 1956/57.

2418: A. van Wieringen and N. Warmoltz: On the permeation of hydrogen and helium in single crystal silicon and germanium at elevated temperatures (Physica **22**, 849-865, 1956, No. 10).

A mass spectrometer examination of the permeability of the elements silicon and germanium to the gases hydrogen and helium has been carried out in the temperature range 967-1207 °C for silicon and 766-930 °C for germanium. Using certain crystal growing and cutting techniques, two kinds of diffusion cells were made by each of which it was possible to determine both diffusion coefficient and solubility as well as the activation energies of diffusion and solution from non-steady-state permeation measurements. No permeation of the gases neon, argon and nitrogen could be detected. It seems that hydrogen in silicon can occur in a readily-diffusible form and also in a state in which it is less mobile. The readily-diffusible form consists of atoms or protons.

2419: N. W. H. Addink: Spectrochemical analysis by means of the D.C. carbon arc (Appl. Spectroscopy **10**, 128-137, 1956, No. 3).

A complete description of the constant temperature arc method of quantitative analysis which has been developed in Eindhoven is given, with tables of the empirically determined K-values reported so that they can be checked in other laboratories. The method consists of completely volatilizing 5 mg of a powdered sample in a shallow anodic crater of a carbon arc, with the addition of materials to modify the rate of volatilization if required. The line intensities are calibrated and corrections are made by comparison with selected Fe lines, originating from a "standard light source" so as to get comparable analytical results; the calculations required are illustrated by several examples, which indicate the relative accuracy of the method to be approximately 10%.

2420: K. Compaan and Y. Haven: Correlation factors for diffusion in solids (Trans. Faraday Soc. **52**, 786-801, June 1956, No. 6).

The relation of Einstein $D = BkT/e$, relating the diffusivity and mobility of particles (ions or atoms), must be modified for the case of diffusion in solids, if a vacancy mechanism holds, because there will be a correlation between successive steps of a particle, even if the steps of the vacancies themselves are uncorrelated. Certain symmetry conditions being fulfilled, the relation of Einstein must be modified by a correlation factor

$$f = (1 + \overline{\cos \vartheta_{i,i+1}})/(1 - \overline{\cos \vartheta_{i,i+1}}), \qquad (1)$$

where $\vartheta_{i,i+1}$ is the angle between two successive steps of a particle. The diffusion problem can be translated into the theory of electrical networks and with the help of measurements in a resistor network the value of $\overline{\cos \vartheta_{i,i+1}}$ is easily evaluated. This has been done for several types of lattices. Other cases, where eq. (1) does not hold, because of lack of symmetry, are treated by similar methods. Appropriate substitutions having been made, the diffusion of associated pairs has been treated as the diffusion of a single particle in a so-called transposed lattice.

2421: C. M. van der Burgt: Les transducteurs piézomagnétiques à noyau massif de ferrite (Commun. Congrès international sur les traitements par les ultra-sons, Marseille 23-28 May 1955; published 1956).

Interim report on some of the results of an investigation into the use of ferroxcube materials for piezomagnetic vibrators. A fuller account has meanwhile appeared in this Review: Philips tech. Rev. **18**, 285-298, 1956/57, (No. 10).

2422: R. Vermeulen: Stereo reverberation (IRE Transactions on Audio, **AU-4**, 98-105, 1956, No. 4).

Reproduction of the article published in Philips tech. Rev. **17**, 258-266, 1955/56.

2423: H. J. Oskam: High-frequency gas-discharge breakdown in neon-argon mixtures (J. appl. Phys. **27**, 848-853, 1956, No. 8).

Breakdown electric fields in a waveguide at a frequency of 9500 Mc/s are presented for a number of neon-argon mixtures at various pressures. The argon percentage is found to have a large influence on the breakdown electric field, just as Penning found for the D.C. discharge. For each argon concentration, only one minimum is found in the curves giving breakdown field as a function of pressure. This result contrasts with the two minima found in the D.C. discharge for some neon-argon mixtures. This can be explained by the difference between the efficiency of energy transfer from the electric field to the electrons in the D.C. case and the high-frequency case. The relation between the concentration of argon and the breakdown fields is discussed. The influence of electric field distortions and the standing-wave ratio in the waveguide are investigated.

2424: W. L. Wanmaker: Contribution à la chimie des halophosphates de calcium (J. Phys. Radium **17**, 636-640, 1956, No. 8-9).

Detailed study of the secondary reactions during

the preparation of calcium halophosphates activated by Sb^{3+} or by Sb^{3+} and Mn^{2+}, in order to suppress the formation of Sb^{5+} or Mn^{3+} ions, or of free Sb_2O_3 and Mn_3O_4, the effect of which is to lower the quantum efficiency.

2425: J. L. Ouweltjes: Quelques considérations sur la transmission d'énergie dans les halophosphates (J. Phys. Radium **17**, 641-644, 1956, No. 8-9).

Measurements have been made of the brightness of halophosphates activated by Sb and Mn, and of the killer effect of iron. The energy transfers between activator, sensitizer and killer ions are discussed. Possibly the greatest part of the killer effect of iron is a mere absorption.

2426: W. Ch. van Geel: Sur la luminescence de couches d'oxydes formées par oxydation électrolytique (J. Phys. Radium **17**, 714-717, 1956, No. 8-9).

Measurements of the luminescence L produced during the formation of layers obtained by anodic oxidation give $L = aI(e^{bd} - 1)$; I current density, d thickness of the layer. When A.C. tension is applied, light flashes are observed at each change of polarity.

2427: P. Zalm: Sur l'électroluminescence du sulfure de zinc (effet Destriau) (J. Phys. Radium **17**, 777-782, 1956, No. 8-9).

A discussion is given of the voltage and temperature dependence of the emittance of an electroluminescent phosphor. The emittance (H) — voltage (V) relation is given by $H = H_0 \exp(-b/V^{1/2})$. The variation of the emittance with temperature depends on the relation between the local field in the phosphor particles and the applied field, the temperature dependence being described by the variation of b with temperature.

2428: J. H. Stuy: Studies on the mechanism of radiation inactivation of micro-organisms, II. Photoreactivation of some bacilli and of the spores of two Bacillus cereus strains (Biochim. biophys. Acta **22**, 238-240, 1956, No. 2).

In order to find out whether bacilli in general could be reactivated by light after ultraviolet inactivation, several different species have been investigated. About half of them did not show any photoreactivation (PHR) while some of them showed a moderate PHR. Only two B. cereus strains could be photoreactivated very easily. The conclusion is drawn that PHR is not so generally occurring among the bacilli, but under entirely different conditions there might be some reactivation. The spores of the B. cereus strains mentioned above did not show PHR. This was not due to their lack of water, since "germinated spores" behaved similarly in this respect. Germination of the spores was carried out in a synthetic medium which permitted a transformation of the spores into a following stage only. During this transformation solids are exchanged for water from the medium, causing a considerable increase of the water content.

2429: J. H. Stuy: Studies on the mechanism of radiation inactivation of micro-organisms, III. Inactivation of germinating spores of Bacillus cereus (Biochim. biophys. Acta **22**, 241-246, 1956, No. 2).

Upon incubation in a germination medium, resting spores of Bacillus cereus lost their resistance against heat, X-rays and ultraviolet radiation. These phenomena were studied in detail using a synthetic germination medium which permitted a transformation of the spores into a following stage only. The results showed that the loss in resistance against heating at 70 °C for 10 minutes was very rapid. The loss of resistance against X-rays followed roughly the same rate. The UV-resistance of the spores, however, increased considerably during the first minute of incubation; thereafter it rapidly fell off. Apart from the UV-increase in the first minute, the simultaneous loss of the three resistances studied suggests that this loss is due to one mechanism. In this respect the water uptake by the spores is considered.

2430: H. B. Haanstra: Einige Bemerkungen über die Polystyrol-SiO-Abdruckmethode (Rev. universelle Mines **99**, 481-485, 1956, No. 10). (Remarks on the polystyrol-SiO replica technique; in German.)

The examination of electron micrographs obtained using the polystyrene-SiO technique shows that the thickness of the SiO layer is not uniform. Using this idea, the author explains that certain zones of the replica which were not perpendicular to the direction of the beam, will appear much clearer than others in transmission.

2431: J. B. de Boer and W. Morass: Berechnung der Sehweite aus der Lichtverteilung von Automobilscheinwerfern (Lichttechnik **8**, 433-437, 1956, No. 10). (Calculation of the visibility range from the light distribution of car headlamps; in German.)

V. J. Jehu has derived a partly experimental,

partly theoretical method for calculating the visibility range in the light of car headlamps of which the light distribution is known. Allowance is made for dazzle by oncoming cars. The results, however, are not always in agreement with observation. The authors have therefore made extensive new measurements of visibility under road conditions. With the help of the results the visibility can be derived graphically. Comparisons between the quality of car headlamps can then be made very simply, knowing their light distribution.

R 310: C. M. van der Burgt: Controlled crystal anisotropy and controlled temperature dependence of the permeability and elasticity of various cobalt-substituted ferrites (Philips Res. Rep. **12**, 97-122, 1957, No. 2).

Dynamic elasticity at remanence, piezomagnetic coupling at remanence, and initial permeability of polycrystalline toroids of various cobalt-substituted ferrites are determined in the temperature range from -196 °C to $+120$ °C (and higher). The compositions are represented approximately by $(M_{1-y}Zn_y)_{1-x}Co_xFe_2O_4$, where M stands for one of the divalent metal ions ($Li_{\frac{1}{2}}^{1+}Fe_{\frac{1}{2}}^{3+}$), Ni^{2+}, or Mn^{2+}. The first-order magnetocrystalline anisotropy constant K_1 of cobalt ferrite is known to be large and positive (2 to 3×10^5 J/m^3). Consequently the incorporation of a small amount of cobalt ferrite in solid solution in other ferrites that have $K_1 < 0$ (ferrous, nickel, manganese ferrites, and presumably lithium ferrites) leads to a compensation of crystal anisotropy at a transition temperature T_0 depending on the amount of cobalt ferrite. At room temperature the compensating effect of cobalt substitutions in ferrous ferrite is nearly 4 times as strong as expected from simple linear interpolation. On the other hand cobalt substitutions in manganese ferrite appear to be less effective than expected, the compensating amount of cobalt being about 4 times that expected. Nevertheless the curves T_0 vs. x for all systems of cobalt-substituted mixed ferrites have a similar shape. The fact that the crystal anisotropy passes through zero at a temperature T_0 implies that around this temperature only strain anisotropy and pore-shape anisotropy remain, so that there exists a small temperature range where the substance is magnetically and magnetoelastically soft. The resulting peaks in the permeability and compliance, together with the overall increase of these quantities with temperature, lead to temperature ranges some-where above T_0 where the permeability and the elasticity are substantially temperature-independent.

R 311: K. S. Knol: A thermal noise standard for microwaves (Philips Res. Rep. **12**, 123-126, 1957, No. 2).

A thermal noise source for the 3-cm waveband, consisting of a heated platinum waveguide terminated with a ceramic wedge, is described. The temperature of the wedge is fixed at the melting-point of gold. With this noise source as a standard, the noise temperature of the Philips noise source K50A is determined to be 21700 °K, with an accuracy of about 5 per cent.

R 312: J. van den Boomgaard and K. Schol: The P-T-x phase diagrams of the systems In-As, Ga-As and In-P (Philips Res. Rep. **12**, 127-140, 1957, No. 2).

For the compounds In-As, Ga-As and In-P the phase relations solid-liquid-vapour have been determined. In-As has a maximum melting-point of 943 ± 3 °C at an arsenic pressure of 0.33 atm. Ga-As has a maximum melting-point of 1237 ± 3 °C at an arsenic pressure of 0.9 atm. For In-P the maximum melting-point is estimated to lie at 1062 ± 7° C at a phosphorus pressure of approximately 60 atm.

R 313: J. L. H. Jonker and Z. van Gelder: The internal resistance of a radio-frequency pentode (Philips Res. Rep. **12**, 141-175, 1957, No. 2).

Besides the direct electrostatic influence of the anode potential upon the cathode current, which is small in radio-frequency pentodes, the authors investigate the effects causing the current distribution between screen grid and anode due to reflection of electrons by the suppressor-grid wires and by the anode (secondary emission). The extra space charge in the cathode space, due to these reflected electrons, is taken into account. The reflected electrons are also distributed between anode and screen grid, as a result of the same effects. A convergent series therefore originates for the anode current, the derivative d_aI/dV_a of which gives the reciprocal value of the internal resistance. To obtain a high value of the internal resistance, the reflection coefficient of the suppressor grid must be small, which can be obtained by taking a very small value for the ratio between the wire diameter and the pitch of this grid. The effective potential of the suppressor grid, however, must be kept low. The calculated values of the internal resistance of radio-frequency pentodes are about 40% higher than the measured ones, which, in view of the large number of effects playing a part, must be considered as a satisfactory result.